

Evaluation of the Census 2000 Quality Assurance Philosophy and Approach Used in the Address List Development and Enumeration Operations

FINAL REPORT

This evaluation reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning.

Authors:

David Morganstein

David Marker

Daniel Levine

Westat, Inc.

Broderick E. Oliver, Project Manager
Decennial Statistical Studies Division

CONTENTS

	EXECUTIVE SUMMARY	iii
1	BACKGROUND	1
2	METHODOLOGY	3
3	LIMITS.....	5
4	HISTORICAL PERSPECTIVE	7
	4.1 1960 Census	7
	4.2 1970 and 1980	8
	4.3 1990.....	9
	4.4 Census 2000	10
5	ADDRESS LIST DEVELOPMENT AND ENUMERATION OPERATIONS	12
	5.1 Background	12
	5.2 Block Canvassing.....	12
	5.3 LUCA 98 Field Verification (Local Update of Census Addresses 1998)	13
	5.4 LUCA 99 Field Verification (Local Update of Census Addresses 1999)	14
	5.5 Update/Leave (U/L)	15
	5.6 List/Enumerate (L/E).....	16
	5.7 Update/Enumerate (U/E).....	17
	5.8 Nonresponse Followup (NRFU)	18
	5.9 Coverage Improvement Followup (CIFU)	18
	5.10 Summary	19
6	THE DEMING PHILOSOPHY.....	20
	6.1 Unique Aspects of the Census.....	20
	6.2 Summarizing Deming	21
7	A SUMMARY OF VIEWS ON QUALITY ASSURANCE	24
8	QUALITY ASSURANCE IN CENSUS ACTIVITIES IN OTHER COUNTRIES.....	30
	8.1 Office for National Statistics, United Kingdom	30
	8.2 Statistics Canada	34
	8.3 Australian Bureau of Statistics (ABS).....	39
	8.4 Summary	41

CONTENTS (CONTINUED)

9	QUALITY ASSURANCE IN CENSUS 2000: STRENGTHS AND WEAKNESSES.....	44
	9.1 Strengths.....	45
	9.2 Weaknesses	46
10	PLANNING QUALITY ASSURANCE FOR CENSUS 2010.....	50
	10.1 Rectifying Major Shortcomings	50
	10.2 Other Suggestions for Consideration	56
	REFERENCES.....	59
	RESPONSE TO WESTAT FINDINGS BY DECENNIAL STATISTICAL STUDIES AND FIELD DIVISIONS.....	62
	APPENDICES	
A	CONTRIBUTING PARTICIPANTS.....	65
B	DEMING’S 14 MANAGEMENT POINTS.....	67

EXECUTIVE SUMMARY

Census 2000 involved numerous field operations and, at its peak, employed almost a half-million temporary workers spread throughout every village, town, county, and state in the United States. Managing the quality of the data produced by this large, decentralized, and transient workforce was a major challenge, which the Census Bureau attempted to meet by introducing an extensive quality assurance program into its ongoing operations. This quality assurance mission had three objectives:

1. To minimize significant performance errors;
2. To prevent the clustering of significant performance errors; and
3. To promote continuous improvement.

This report, prepared by Westat, evaluates the effectiveness of the quality assurance programs developed for and implemented in the major field operations in Census 2000 and, specifically, those operations used to update the Bureau's nationwide address list and directly enumerate the population. Further, it identifies strengths and major deficiencies, and provides a critique of the Bureau's quality assurance philosophy. Finally, it offers recommendations for dealing with the identified problem areas and strategies for improvement that are workable for a large and diverse workforce, recognizing both existing technology and possible future developments.

The Report was accomplished through a combination of approaches, including a review of materials prepared for use in training and in collecting data, examining reports prepared by staff who observed different operations, interviewing a range of Census Bureau staff who had been closely involved in the many phases of the quality assurance program (and were still employed at the Bureau), and by examining such materials as were available concerning the operations, as well as the successes or failures, of the quality assurance programs. Unfortunately, such data are seriously limited, since the Census Bureau is well behind schedule in completing its series of evaluation profiles on its various quality assurance programs. We also reviewed Deming's philosophy, particularly as it might be applied to an effort as extensive and short-lived as a decennial census. As a final step, Westat held discussions with staff in the statistical offices of the United Kingdom, Canada, and Australia, all of which had recently conducted a national Census, in order to determine how these countries approached the application of quality assurance in their Census programs, and the possible relevance of their actions to future quality assurance planning of the Census Bureau.

It should be noted that, although our focus was on a number of designated field operations, it quickly became apparent that, whatever their responsibilities in Census 2000, staff viewed our efforts as providing a much broader forum for expressing their thoughts on the overall quality assurance program, as carried out during the collection period. We also would emphasize that, in addition to fact, this Report reflects the opinions, attitudes, and perceptions of the staff who provided their views. However, to the extent that these views are widely shared or represent the view of a person in a position of significant authority or responsibility, they provide important insights into staff attitudes towards quality assurance and deserve attention.

The extensive range of reactions to the Bureau's quality assurance programs and to the application of quality assurance in general is found in the Report. At this point, we would emphasize the "strengths" and "weaknesses" of the Bureau's quality assurance planning, organization, implementation, and management.

On the positive side:

- Consistent with its mission statement, Census 2000 continued the tradition, initiated in the 1960 Census, of incorporating into Census 2000 Field operations numerous activities described as quality assurance. This commitment to quality and quality assurance, demonstrated in five censuses over a 40-year period, certainly is a significant "strength". The Bureau also met its objective for quality assurance that it be completely transparent in Census 2000. To that end, materials used to train enumerators and first level supervisors contained specific references as to why quality assurance was important and as to how it would be implemented, and all enumerators were exposed to the concept of and need for "quality" performance and, generally, were measured against the established standards. Finally most operations, unlike the situation in the 1990 Census, had some form of quality assurance process in place;
- Given these many developments, it is not surprising to find that the overall perception throughout the Bureau, and at all levels, is that the Census 2000 quality assurance Field program was an important element in preventing significant errors, and in preventing the clustering of significant errors. Although errors of both types did occur, for the most part, they were caught expeditiously and rectified. On this basis, the quality assurance Field programs can be viewed, generally, as successfully meeting the first two elements of the Bureau's quality assurance mission to prevent significant errors and to prevent the clustering of significant errors; and
- As to the Bureau's goal of "making the best use of the available technology and statistical process tools with respect to its desire to promote timely and continuous improvement throughout the field operations", we conclude that the answer is mixed. Based on the Bureau's evaluations to date and comments from those involved, many of the Bureau's early activities in preparing for Census 2000 are seen as having utilized a full quality assurance approach that met the Bureau's stated goal of promoting timely and continuous improvement. Examples of activities considered as

having successfully utilized a quality assurance approach include preparing the geographic framework, and printing questionnaires and related forms. However, in the context of what actually transpired during the data collection phase, the perception is less clear and decidedly mixed, as noted in the Report.

It should not come as any surprise that an activity as vast and diverse as the decennial census would exhibit some “weaknesses” or faults. And it should be a source of satisfaction that the Bureau actively seeks to locate, learn about, and understand such faults, in order to improve the upcoming decennial census. To that end, the report provides a broad view of weaknesses identified in the quality assurance program implemented for the data collection phase, reflecting in some cases factual evidence and, in others, the views and opinions provided us. At this point, however, we highlight some of the major problems. Specifically:

- The lack of a senior management team throughout the decennial effort with responsibility for coordinating and approving the overall quality assurance plan and reviewing the implementation;
- The failure to ensure the independence and importance of quality assurance throughout the organization. Briefly stated, quality assurance was not perceived as an equal partner, nor was quality assurance staff given either the necessary authority or the required freedom to complete its task successfully and, finally, dedicated quality assurance staff was not assigned to Regional Census Centers;
- A vital aspect of the quality assurance program—especially for promoting continuous improvement—real-time capture and dissemination of data during the data collection process, with which to monitor, evaluate, and react, was not implemented; and
- The apparent low priority afforded the quality assurance effort throughout the entire decade. As illustration, it appears that only very limited quality assurance research was undertaken during the intercensal period, that quality assurance staff from the 1990 decennial effort were not retained, that relatively limited resources were made available to plan and develop the 2000 quality assurance program until well into the decade, and that, currently, some two years after completion of data collection, relatively little factual information is available on the effectiveness of the Census 2000 quality assurance programs.

Section 10 contains our recommendations towards planning quality assurance for Census 2010. Some of the key suggestions are noted below:

- Ensure that the quality assurance effort is seen as an integral and important element in the Census 2010 program. To that end, provide adequate resources, both in funding and staff, and initiate early planning, research, and testing for the quality assurance phase;

- Involve the executive staff in supporting and monitoring quality assurance efforts, especially throughout the data collection phase;
- Establish a senior management team to coordinate and approve the overall quality assurance plan and, throughout the decennial period, to review progress and resolve issues;
- Establish the equality of quality assurance relative to production. Simply put, quality assurance must be seen, understood, and accepted as an essential element of the Census and as an equal partner at all levels of planning, implementation, and review;
- Develop and implement a Management Information System component which provides management, in real-time, with relevant information on the quality of the data collection elements; and
- Expedite the documentation of Census 2000 and establish ready access to the information. Some examples include the quality assurance evaluation program and the documentation of the experiences, problems, and solutions, suggestions and recommendations of staff, and the accumulation of memoranda detailing problems, issues, and solutions.

1. BACKGROUND

Census 2000 involved numerous field operations and, at its peak, employed almost a half-million temporary workers spread throughout every county in the United States. Managing the quality of the data produced by this large, decentralized, and transient workforce was a major challenge, which the Census Bureau attempted to meet by introducing an extensive Quality Assurance (QA) program into its ongoing operations. This QA mission had three objectives:

1. To minimize significant performance errors;
2. To prevent the clustering of significant performance errors; and
3. To promote continuous improvement.

In addition to providing a myriad of important data, each decennial census contains an important evaluation component through which the Census Bureau attempts to evaluate the effectiveness of its programs, both in order to provide users with some indication of the quality of the census and the results, and to understand how well its programs succeeded. This latter effort identifies deficiencies and problems, and its findings are used to improve future decennial programs and activities. In this connection, the Census Bureau asked Westat to assist in studying the QA process utilized in a number of specific, key field collection activities (described in Section 5), and to identify strengths and major deficiencies. Westat also was charged with providing a critique of the Bureau's QA philosophy and offering creative solutions and recommending strategies for improvement that are workable for a large and diverse workforce, recognizing both existing technology and possible future developments.

The purpose of this report is to evaluate the effectiveness of the QA programs developed for and implemented in the major field operation in Census 2000 and, specifically, those operations used to update the Bureau's nationwide address list and directly enumerate the population.

The Report begins with brief discussions of the Methodology (Section 2) and the Limits of the Report (Section 3), which are followed, in Section 4, by a brief history of QA efforts in recent decennial censuses. Section 5 describes both the eight field operations covered by this study and the relevant QA activities implemented for each operation. Section 6 summarizes Deming's philosophy concerning quality assurance and compares the Census Bureau's approach with Deming's criteria. In Section 7, we present a summary of views, reactions, and opinions collected from selected staff that was involved with different aspects of the QA program. Section 8 looks at how some selected foreign countries evaluate the effectiveness of their QA activities in census taking, and attempts to relate their

approaches to the U.S. effort. Section 9 summarizes the results of the previous chapters, addresses the specific questions raised by the Census Bureau, and describes the strengths and weaknesses of the Bureau's efforts. Finally, Section 10 looks ahead; to this end it suggests approaches and developments to be considered in the planning of QA for Census 2010.

2. METHODOLOGY

This Report was accomplished through a combination of approaches, including a review of written materials prepared for use in training and in collecting data, by examining reports prepared by staff who observed the operations, by interviewing key Census Bureau personnel who had been closely involved in the many phases of the QA program, and by examining such materials as were available concerning the operations of the QA programs. Although our initial focus was on the eight major field operations described in Section 5, the review became more broadly focused on the overall field QA program, especially as it was carried out during the decennial collection period.

The report also reflects our review of the training and operational materials prepared by the Bureau for each of the Field operations, a range of Travel Reports prepared by staff who visited local census offices (LCOs) and observed the operations during their implementation, and such other materials made available to us that reflected on the conduct of the QA effort. We would note, however, the scarcity of information or reports, including from the Management Information System (MIS) system, on the successes or failures of the QA programs, or their effect on the operations.

As a first step, Westat reviewed Census Bureau documentation concerning QA in Census 2000, starting with Census 2000 Operational Plans describing the proposed QA program, followed by the large and diverse body of materials used in the training of supervisory staff and enumerators for each of the specified field activities, and concluding with local office materials pertaining to QA, such as preparing for field QA and editing and handling the resultant QA forms.¹ Subsequently, meetings were held with staff throughout the Census Bureau that had been responsible for or actively involved in the different aspects of the QA program, ranging from policy decision making through the development of specific QA programs and the preparation of materials, as well as staff who had served in field offices with responsibility for carrying out the QA program and accomplishing the collection of data. Westat next reviewed the substantial body of material concerning QA in selected field operations which had been obtained through a range of debriefing efforts, as well as field trip reports prepared by Washington QA staff who had visited LCOs during the conduct of the census and reported their observations of different operations, including the implementation of the QA programs. The Census Bureau is in process of preparing a series of evaluation profiles on the various QA programs but, unfortunately, these were not yet available in time for this review. As a final step, Westat undertook discussions with staff in the statistical offices of the United Kingdom, Canada, and Australia, all of which had recently conducted a

¹ See references for a listing of the materials reviewed.

national Census. The purpose of these conversations was to determine how these countries approached the application of QA in their Census programs, and the possible relevance of their actions to the future QA planning of the Census Bureau. The results of these activities are reflected in the report.

Although our initial focus was on the eight field operations described in Section 5, it quickly became apparent that, whatever their responsibilities in Census 2000, staff with whom we came in contact viewed our efforts as providing a much broader forum for expressing their thoughts on the overall QA program as carried out during the collection period. In addition, we found that relatively little or no data concerning the usefulness and success of the individual QA programs were currently available, although these data are in process of being compiled. For these reasons, this report is directed more broadly towards an assessment of QA in the overall data collection process for Census 2000.

3. LIMITS

As noted, this report reflects the opinions, attitudes, and perceptions of the staff who shared their views with us. It should be emphasized that some of these comments reflect “perception,” whereas factual reality, indeed, may be somewhat different. However, to the extent that these views are widely shared or represent the view of a person in a position of significant authority or responsibility, they provide important insights into the attitudes towards quality assurance, and should not be dismissed summarily. Our discussions evoked both positive and negative comments, sometimes on the same issue or program, and even from the same discussant. We also would note our surprise at the extent and relative consistency of the negative reactions and comments, or the dichotomy of view, concerning the QA effort, forthcoming from those with whom we spoke. Although positive comments about the QA efforts were expressed, and are reflected in this section, we believe that most participants were identifying areas requiring improvement and that their comments addressed that view.

Given the elapsed time since the completion of the data collection effort, it was not possible to draw any sort of “representative” sample of those who participated in the Field. For example, neither Regional Office Directors nor Assistant Regional Census Managers (ARCMs), whose knowledge on these issues would have been useful, were interviewed for this study. Our investigation, of necessity, was limited to those who had been closely involved in different aspects of the QA operation and were still members of the Bureau’s current staff. We did make every effort, however, to locate persons at every level of the operation and whose responsibilities varied widely. To that extent, we were successful, talking with staff responsible for both the early phases of the planning and for the later stages of preparing materials and implementing the planning; we spoke with staff who were involved at the LCO level and those who were involved at the Headquarters level. We also were fortunate in that members of the Executive Staff who had the decision making responsibility at different stages and phases of the decennial effort shared their thoughts, views, and opinions as to the successes and failures of the QA effort, as perceived by them.

Since our discussions were conducted many months, even years, after the completion of the Census, certain issues may have been overlooked or remembered differently than actually experienced, or even viewed in a different light than if the discussions had occurred during or immediately after the completion of the specific activity.

Finally, although an outline of the topics and issues was developed to guide our discussions, of necessity our time with each participant was limited. Nonetheless, we believe that sufficient information was gathered to support both the objectives of the study and the conclusions derived.

4. HISTORICAL PERSPECTIVE

Before addressing the system implemented by the Bureau for Census 2000, it may be helpful to provide some historical perspective, by reviewing briefly the Bureau's approach to quality in several recent decennial censuses.

4.1 1960 Census²

The 1960 census was the first U.S. census to use "statistical quality control" for field operations, that is, a statistical system which required specific action for varying levels of error, and which applied at all levels of the operation. Thus, information collected from households by enumerators and entered into enumeration books was first reviewed by crew leaders; second, quality control checks were carried out in the local office, consisting of inspecting completed enumeration books received from the field to determine whether or not the crew leaders were doing an adequate job of inspection. Each of the temporary field office District Supervisors was assisted by one or more Technical Officers who, in addition to their major role in training crew leaders on the technical content of the field enumeration, were responsible for supervising the quality control program. Finally, Program Technicians from the Regional Offices visited each local office periodically throughout the census period to review evaluation forms prepared by Technical Officers in the review of Crew Leaders, and to review the various quality control reports received by the Technical Officers from Crew Leaders and field reviewers. In turn, the Program Technicians made formal reports of their observations to the Regional Field Directors, using evaluation forms similar to those used by Technical Officers in evaluating Crew Leader activities. They also provided assistance and guidance to the Technical Officers. The existence of a formal quality control system also was seen as having a significant intangible effect of creating a climate that helped engender good quality. It was believed that the enumerator, knowing that his work was to be checked, was likely to use more care than might have been the case otherwise, to ensure that work would pass inspection. Similar effects were expected throughout the process.

The most direct evidence of the effectiveness of quality control in the 1960 census is that some 1,400 enumerators were eliminated early in the process, as a result of quality control inspection of their work. Since this group produced a disproportionate share of the total errors, its release was seen as having had a great impact on the quality of the census. Retraining of enumerators who were not released also affected the quality. For example, the number of reported errors was cut in half between first review

² U.S. Bureau of the Census, *Quality Control of the Field Enumeration*, Census of Population and Housing: 1960. Washington, D.C. 1967.

and final review of the enumeration books. Another indicator of the positive impact of the quality control system is that about 800 assignments in each stage of the census were rejected by crew leaders and reassigned to other enumerators for cleanup, as were substantial numbers of enumeration books in each of the census stages. A major finding was that a small proportion of the staff was responsible for a large proportion of the errors.

The Bureau's evaluation of its Quality Control (QC) Program concluded that "the quality controls, although reasonably effective, fell considerably short of their potential." For example, estimates developed prior to the initiation of the field work of the expected number of enumerators who should have been released as a result of first review, the expected number of assignments to be rejected on final review, and the expected number of enumeration books to be rejected in office review, were far in excess of what actually was recorded during the Field operation. Studies carried out after the conclusion of the census showed the original estimates to be virtually correct, reflecting serious problems in the implementation and errors on the part of those conducting the review. Further, QC staff spent an inordinate amount of time on non-QC activities.

Nonetheless, the evaluation went on to note that, "nevertheless, statistical quality control in 1960 was generally regarded as making important contributions, both in terms of improvement over what took place in earlier censuses and in its own right. It was a marked advance over past efforts, primarily because formal specifications were provided for the various crew leader and office activities."

4.2 1970³ and 1980⁴

Both the 1970 and 1980 Censuses, for the most part, relied on the "inspection and repair" method of QC and, in large measure, replicated the program established in 1960, with some refinements. Assessment of the 1980 program suggested that this approach had not been completely successful and, in part, was ascribed to the fact that production and quality responsibilities resided in different management areas. Responsibility for production rested with Field, whereas responsibility for quality resided elsewhere in the organization. When management inquired about progress, the response was perceived in terms of quantity, rather than quality, of work, which was translated into a perceived priority on the production side within the organization's structure. The result was a decided adversarial relationship between the component entities.

³ U.S. Bureau of the Census. U.S. Census of Population and Housing: 1970. Procedural history PHC @-1. 1976.

⁴ Information on the experience in the 1980 Census is found in Chapter 1, Introduction and Background, U.S. Bureau of the Census, Effectiveness of Quality Assurance: 1990, Report Series CPH-E-2. Washington, D.C. 1993.

4.3 1990⁵

In planning the QA program for the 1990 Census, the Census Bureau adopted the Deming philosophy, with its approach toward total quality improvement, for the decennial census program. Deming argued that quality results from the prevention of defectives through process improvement, not solely inspection. Inspection judges the quality of finished products; defective items are scrapped or reworked. The goal of process improvement is to build products correctly the first time, and to continuously reduce the variation of the results around the desired outcome. Statistical process control plays a major role in achieving this objective, as does management involvement and commitment to the quality improvement process. Four major components were identified:

1. Build quality into the system;
2. Constantly improve the system;
3. Integrate responsibility for quality with production; and
4. Clearly differentiate between QA (the prevention of error through process improvement) and QC (measuring error after the fact and attempting to make corrections).

The 1990 census program dealt with the “responsibility” problem by assigning the production side (the Field) responsibility for quality. With this added responsibility, not only did the job have to get done, the job, now, had to be done well. This change was one of the most difficult to implement during the 1990 census. Traditionally, Field staff, the “production side,” devoted all their energies to ensuring that the census was conducted in an efficient and timely manner. Many found it difficult to take on the additional task of “quality.” Although substantial efforts were invested in attempting this change, it was not successful. Discussion with senior staff at the time of the 1990 Census indicated that the shift in responsibility to Field for the QA program failed to resolve the problem for the same reasons as perceived earlier; namely, that Field management did not provide sufficient support to counter the emphasis on production.

To accomplish the stated goals, the Bureau also attempted to simplify its manual records and summaries, and to develop software to support the quick capture and transmittal of data quality. Efforts were made to measure performance both during training and during production. Timely feedback was deemed essential and built into all levels of the activity. It was hoped that the increased use of automation

⁵ U.S. Bureau of the Census, Effectiveness of Quality Assurance, 1990 Census of Population and Housing, Evaluation and Research Reports, Series CPH-E-2. Washington, D.C. 1993.

would make it possible to extend the use of QA to new areas and to make the results available more quickly and efficiently. The Bureau also implemented a range of actions designed to result in an effective quality assurance program. Examples include establishing working groups and quality circles to effect improved communication, reducing the ratio of enumerators to supervisors to allow supervisors more time for reviewing enumerators' work, feedback of information, and counseling and retraining, providing more and better education and training of the staff, especially training on the job and, finally, instituting a system to accurately measure performance, document the characteristics of the errors, and provide relevant and sufficient information to management so that feedback could be given. To assist in the accomplishment of the QA effort, quality assurance technicians were assigned to each of Regional Census Centers (RCCs), to enhance local management's awareness of QA objectives and importance, assist in monitoring adherence to the QA requirements, and identify problems and refer them to RCCs and District Office management. This program was seen to have accomplished all three of its objectives, in general, although it was noted that the QA technicians might have been considerably more effective had they not been assigned part-time, for the most part, to this activity, as well as required to undertake additional responsibilities beyond QA, such as recruiting, training, observation, reinterview, and some office responsibilities.

The techniques employed to measure field performance included the use of pre-operational sampling (for pre-listing operations), concurrent monitoring (for Update/Leave), sample suppression (for Precanvass), and Reinterview (for nonresponse). In general, the Bureau concluded that its QA approach and programs was effective overall, although its evaluation provided numerous suggestions for further examination and improvement.

4.4 Census 2000

Deming's quality assurance philosophy was reinforced in Census 2000. To implement Deming's approach, an extensive QA program was proposed. For example, during training, workers were tested on their knowledge and given practice fieldwork. Once the census was underway, supervisors measured performance and gave the staff precise and timely feedback. Other QA activities during the data collection phases included initial observations of the staff, informal reviews of performance, dependent reviews of completed work, reinterviews of respondents, and office reviews. Various statistical process tools were employed, including check sheets to tally problem areas by frequency of type, and a modified control chart to detect potential enumerator fabrication. Acceptance sampling also was included, not as a direct form of quality control, however, but rather as an audit tool to ensure, first, that the completed product conformed to the Bureau's quality requirements and, second, that significant performance errors did not impact communities and small geographic entities disproportionately. For the most part, however,

the techniques that were actually employed in Census 2000 during the collection phase were largely the same as those utilized in the immediately preceding censuses, with some modification to improve effectiveness.

5. ADDRESS LIST DEVELOPMENT AND ENUMERATION OPERATIONS

5.1 Background

The primary goal of Census 2000 was to determine how many people reside in the United States, precisely where they live, and their demographic characteristics. Census 2000, the Nation's largest and most complex peacetime mobilization, included numerous critical phases, such as developing an up-to-date nationwide address list, printing questionnaires, delivering questionnaires to households – via mail and in-person, enumerating segments of the population directly, and tabulating the results. To ensure accuracy and completeness of each of these and all other phases, the Bureau instituted QA/QC measures throughout Census 2000.

For this report, the focus is on the QA operations and activities applied to eight major field operations: Block Canvassing, Local Update of Census Addresses (LUCA) 98 Field Verification, LUCA 99 Field Verification, Update/Leave (U/L), List/Enumerate (L/E), Update/Enumerate (U/E), Nonresponse Followup (NRFU), and Coverage Improvement Followup (CIFU). Broadly speaking, these operations served two purposes: Address List Development and Enumeration.

The Address List Development operations (Block Canvassing, LUCA 98 Field Verification, and LUCA 99 Field Verification) were designed to update the Master Address File (MAF), a file that lists most residential addresses in the United States. The Census Bureau linked each living quarter to its unique geographic location through its Topologically Integrated Geographic Encoding and Referencing database (TIGER). A complete and accurate MAF and TIGER are essential elements of a mail census.

The Enumeration operations (U/L, L/E, U/E, NRFU, and CIFU) were designed to enumerate specific segments of the population directly. In the following sections, we describe each of the eight operations mentioned above in more detail, as well as its corresponding QA activities.

5.2 Block Canvassing

The MAF for city-style addresses was created by combining addresses from the Census Bureau's 1990 Census Address Control File with addresses in the United States Postal Service (USPS) Delivery Sequence File. The Block Canvassing operation was one of two field operations implemented in Census 2000 to update the MAF in areas containing city-style addresses (LUCA 98 Field Verification was the other). For Block Canvassing, enumerators canvassed each and every road and street in areas of

city-style addresses looking for every place where people live or could live, and comparing the address of each living quarters with the pre-listed addresses in their Address Binders. Enumerators either verified the pre-listed addresses, or corrected or deleted them. Living quarters that were not pre-listed were added to the Address Binders. The Block Canvassing operation was carried out between January and May of 1999, and included the following quality assurance activities:

- Initial Observation: Soon after each enumerator was given an initial assignment, the Crew Leader observed the enumerator updating the Address Binder and corresponding census maps at approximately 10 addresses or for 2 hours, whichever was less. The Crew Leader tallied any critical listing or mapping errors made by the enumerator and provided the enumerator specific feedback. If the enumerator made too many critical errors, the Crew Leader provided the enumerator on-the-job training. Otherwise, the enumerator was allowed to work independently;
- Weekly Observation: Once the enumerator passed the Initial Observation, he/she was allowed to work independently; however, his/her work was still subjected to the same type of review on a weekly basis, beginning with the second week of work. Similarly, the Crew Leader provided structured feedback and correction when necessary to help the enumerator continually improve performance;
- Dependent QA: This was another weekly quality check; however it was performed on a sample of twenty housing units (HUs) in each Assignment Area (AA). Once again, the Crew Leader tallied the critical listing and mapping errors observed. If the enumerator made too many critical errors, the AA failed this quality check and was recanvassed, that is, another enumerator canvassed the entire AA to verify the Address Binder and map updates and make corrections where necessary; and
- Office Review: This quality check was performed in the LCOs. When the field staff submitted completed Address Binders and census maps to the LCOs, office clerks performed a formal review of these items for completeness and accuracy. If the clerks found mistakes that were correctable in the office, they corrected them. The clerks tallied the errors and determined if the AA passed or failed.

5.3 LUCA 98 Field Verification (Local Update of Census Addresses 1998)

The Census Bureau created partnerships with local and tribal governments to improve the address lists for their jurisdictions. Beginning in July 1999 and continuing through December 1999, cooperating local governments with predominately city-style addresses (house number and street name) reviewed the MAF listings and corresponding census maps of their areas for completeness and accuracy. About half of all eligible communities took advantage of this program. The local officials updated/corrected these items, supplying added housing units, deleting nonexistent housing units, and/or correcting the address/geographic location of housing units. Census enumerators subsequently visited these units to verify the updates. The original plan called for including the LUCA Field Verification addresses in the Block Canvassing operation, but due to delay in implementing the LUCA Field

Verification program, a separate field operation was developed for those LUCA addresses which could not be included in the Block Canvassing operation. The following is a list of the QA activities that were implemented to ensure that the enumerators produced work of an acceptable quality:

- Initial Observation: As soon as possible after the enumerators completed training, the Crew Leaders observed them verify or correct the listings and associated census map for a sample of 10 HUs. The Crew Leader tallied and recorded any critical listing or mapping errors that the enumerator made, for example, a failure to enter the correct action code for the HU. The Crew Leader provided the enumerator feedback and/or retraining, as necessary. As with all of Initial Observations, if an enumerator still continued to have difficulty after retraining, the Crew Leader discussed the problem with his/her supervisor and, if necessary, the enumerator was let go;
- Dependent Verification: Crew Leaders assessed the accuracy of the deleted addresses in an entire AA by checking the accuracy of a random sample of 10 of them (or less). If the Crew Leader found an incorrect delete, the review continued until 10 consecutive correct deletes were found. Mistakes were corrected. Based upon the number of critical errors, the enumerator was allowed to continue to work, received retraining in specific areas, or was released; and
- Office Review: As with the previous operation, LCO clerks performed a formal review of the completed Address Binders and census maps. The clerks checked the entries on these items for legibility, completeness, and consistency between updates in the Address Binder and the corresponding census map. Where possible, the clerks made corrections. Otherwise these items were returned to the field for repair.

5.4 LUCA 99 Field Verification (Local Update of Census Addresses 1999)

This operation was the complement to LUCA 98 Field Verification for places in areas with predominately non city-style addresses. Between May and October of 1999, cooperating local and tribal governments in areas where most of the addresses had no street name and/or house number reviewed the MAF listings and census maps for their respective areas and identified blocks containing potential coverage problems. Census enumerators in the field conducted a dependent canvassing of the identified blocks, using the existing MAF based on the results of the Address Canvassing, and resolved discrepancies found, through updating and correcting the list and through deleting HUs listed in error. The LUCA allowed local officials to improve the accuracy and completeness of the Census address list. The LUCA Field Verification operation was conducted to verify the HUs in areas that were questioned during the LUCA program. Address list review listers canvassed the areas in question to check addresses and determine the correct number of HUs in the area. In order to accomplish this task, the lister:

1. Assigned action codes for every address on the address listing page;
2. Added HUs that were not already on the listing page to the “add” page;

3. Updated the AA map whenever necessary; and
4. Deleted addresses not found in the block.

The following is a list of the QA activities that were implemented to ensure that the enumerators produced work at an acceptable quality:

- Initial Observation: An initial observation on the lister's first Field Assignment (FA) was performed by the field supervisor during the lister's first three days on the job, to ensure that the lister could produce work according to the established procedures. The observation covered ten HUs. During the observation, the field supervisor identified and corrected critical errors and discussed any other errors with the lister. Based on the review, the supervisor determined whether to retain the lister and the need for additional training;
- Daily Reviews: During the daily meeting with enumerators, Crew Leaders informally checked the Address Binders and block maps for legibility, completeness, and consistency. They also compared any remaining addressed questionnaires with the listing pages to be sure that questionnaire were delivered to all HUs, including vacants; and
- Office review of listings and maps.

5.5 Update/Leave (U/L)

This operation, which took place in March, 2000, and involved some 15 percent of all addresses, was conducted primarily in rural areas with predominately noncity-style mailing addresses (i.e., mainly rural route and P.O. Box addresses). Census enumerators delivered questionnaires to the HUs in these areas and updated the corresponding Address Binder and census maps. The QA program for U/L consisted of the following activities:

- Initial Observation: Soon after the enumerators were given their first assignment, the Crew Leaders observed each enumerator interview residents at ten houses, in order to verify/update the address list and block map and leave a questionnaire. The Crew Leaders tallied critical errors and provided the enumerator feedback and/or retraining as necessary;
- Daily Reviews: During the daily meeting with enumerators, Crew Leaders informally checked the Address Binders and block maps for legibility, completeness, and consistency. They also compared any remaining addressed questionnaires with the listing pages to be sure that questionnaire were delivered to all HUs, including vacants;

- Dependent Check: At the completion of each AA, the Crew Leader verified a random sample of twelve HUs. The sample was split between two randomly selected blocks. In each of these blocks, the Crew Leader canvassed the first six consecutive houses on the ground to verify that what was observed was reflected in the Address Binder and on the block maps. The Crew Leader tallied the errors observed and provided advice, criticism, and training as needed. If the enumerator made too many critical errors, the AA failed this quality check and was recanvassed by another enumerator; otherwise it was accepted; and
- Office Review: If office staff identified incorrect, inconsistent, or illegible entries or other unresolved problems during their formal review of the Address Binders and census maps, they corrected them where possible; otherwise, the Binder was returned to the field for repair.

5.6 List/Enumerate (L/E)

In rural areas where residential mail delivery is uncommon (about 1 percent of all addresses), enumerators canvassed their assigned areas to locate and register (list address and spot address on a map) every living quarter and enumerate the household. This operation was conducted between March and May 2000. To ensure the integrity of the data produced, the following QA programs were implemented:

- Formal and Informal Reviews: To ensure that the enumerators were making complete and accurate entries on the questionnaires, in the Address Registers, and on the block maps, Crew Leaders reviewed the enumerators' work and provided feedback;
- Dependent Check: At the completion of each AA, Crew Leaders (or their assistant(s)) checked a random sample of six HUs (two random starts—three consecutive housing units per start) to verify the accuracy of the enumerators' listings and map spots with what was on the ground. Crew Leaders made any necessary corrections and documented their findings. If the enumerator made too many critical errors, the AA failed this quality check and was recanvassed;
- Office Review: Clerks at the LCOs formally reviewed all completed Address Registers and census block maps. If necessary, they were returned to the Crew Leader for repair;
- Reinterview: Each week, a computer-generated report listed the names of enumerators in a given Crew Leader District (CLD) whose questionnaires completed in the previous week differed significantly from the collective results of their co-workers on one or more of the following housing unit characteristics: *number of vacants*, *number of partial interviews*, *number of single-person households*, and *average number of persons per household*. If the Field Operations Supervisor (FOS) could not justify the reason for any of the detected differences, a sample of 7 of the enumerator's subsequent checked-in questionnaires was selected for Reinterview;

- Each designated household in the sample was reinterviewed to determine if the enumerator in question visited the household and enumerated it correctly. Depending upon the results obtained, the questionnaire was accepted as correct or judged to be falsified. The falsified questionnaire and all other questionnaires completed by that enumerator were rejected. This portion of the Reinterview operation was called an *Administrative Reinterview*;
- If the findings from this review were inconclusive, the Reinterview supervisor had the option to place additional questionnaires for that enumerator in Reinterview. This program was called a *Supplemental Reinterview*; and
- Operation Control System (OCS) 2000 Data Entry: Completed questionnaires were sent to the LCOs for data entry into a control system. To ensure the accuracy of the entries for unit status, population count, and vacancy status, an automated edit checked these entries for completeness and consistency.

5.7 Update/Enumerate (U/E)

Conducted between March and May 2000, this rural/urban operation took place in areas containing high numbers of seasonal vacants, American Indian Reservations, and Colonias (about 5 percent of all addresses). Census enumerators updated the addresses and census maps in their assigned areas and enumerated the residents of each listed household. As part of the address update, enumerators verified the address, or corrected, deleted, or added a missing address. The following QA programs were implemented to ensure that the end product met the Census Bureau's standard of quality:

- Formal and Informal Reviews: Crew Leaders performed a thorough check of the questionnaires and a cursory check of the Address Binders, and census maps for legibility, completeness, and consistency during the daily meetings with enumerators. Crew Leaders asked enumerators to fix mistakes in their presence, if possible. Otherwise enumerators returned to the field to make the corrections;
- Dependent Check: When an enumerator finished updating an Address Binder and the corresponding census maps, the Crew Leader performed a formal review of a random sample of six consecutive HUs. The Crew Leader canvassed these six units in a clockwise direction (always making right turns) and compared the units found to the units listed in the binder and spotted on the block map (if in a rural area). The Crew Leader corrected any mistakes found and recorded the findings on a QA form. If the enumerator made too many critical errors, for example, forgot to add a missing unit, the entire AA failed the quality check and was recanvassed;
- Office Review: Each Address Binder and corresponding census maps had to pass this review before being sent to the National Processing Center. Items that could not be corrected in the office were returned to the field for repair;
- Reinterview: Similar to the Reinterview described in the L/E operation. The only difference was that the Administrative Reinterview incorporated the additional variable, "number of deletes"; and

- The OCS 2000 Data Entry: See description in the L/E operation.

5.8 Nonresponse Followup (NRFU)

Beginning shortly after the mail-out phase (during the last week of April), and continuing into early July, a nationwide force of almost 500,000 enumerators visited households that did not return a completed questionnaire to verify/update their address and complete a questionnaire. Enumerators also added any missed units. The QA program for NRFU, the largest of all the field operations consisted of the following programs:

- Formal and Informal Reviews: When enumerators turned in questionnaires to the Crew Leaders, they were reviewed to make sure that they were filled out completely and correctly. Crew Leaders also informally reviewed the Address Binders and census maps (rural areas only);
- Reinterview: The NRFU operation included the same Reinterview program as the L/E and U/E operations, but with two differences: the NRFU Administrative sample contained 10 questionnaires as opposed to 7 for L/E and U/E; and the NRFU Reinterview contained a *Random Reinterview* component that targeted all enumerators during the first three weeks of their assignment. The objective was to identify, as soon as possible, enumerators who were fabricating data or producing shoddy work; and
- The OCS 2000 Data Entry: See description in the List/Enumerate operation.

5.9 Coverage Improvement Followup (CIFU)

This operation followed NRFU and lasted approximately three months. It was conducted in 3 waves, with each wave lasting about a month. The CIFU operation involved follow-up visits to housing units classified as vacants or deletes in NRFU. Residual NRFU cases, such as mail return forms that had been checked in but were lost or blank, also were included, as were newly constructed additions and late additions from U/L and the Delivery Sequence Files. The CIFU enumerator's task was to complete a questionnaire where required for each of the units. For the vacant and delete units from NRFU, CIFU served as an independent verification of the housing unit status on Census Day, April 1, 2000. Further, telephone followup was used to reinterview all households with discrepancies between the reported count of persons in the household and the number of pages completed. The following are the QA activities that we implemented in CIFU:

- Formal Review: Crew Leaders (or their assistants) checked each questionnaire submitted by enumerators to make sure that it was filled out completely and correctly;

- Dependent Verification: When enumerators submitted questionnaires, Crew Leaders visited a sample of the households classified as vacant or delete to verify/correct their unit status and population count; and
- The OCS 2000 Data Entry: See description in the L/E operation.

The QA activities described above did not comprise the totality of the Bureau's quality assurance program; they comprised the major ones. Other QA activities included assignment preparation (in the LCOs) to ensure that the contents of each Address Binder and Map Pouch were complete and correct and that the Address Binder and its corresponding Map Pouch agreed. In addition, QA activities were included in the labeling and distribution of the questionnaires. In line with the Deming philosophy, the Census Bureau built quality into its operation, implementing preventive measures throughout. For example, during training, enumerators were tested and given practice fieldwork. Following practice fieldwork, the trainees engaged in a discussion of their experience and asked questions. These activities improved their understanding of the operation and prevented future mistakes.

5.10 Summary

The range of QA programs reflects the combination of historical precedence and innovation and improvement—that is, many of the programs derive from similar efforts in previous censuses, with modifications and additions reflecting the experience of the most recent census, and with sufficient testing to confirm the value of the changes. In addition, the programs were extended to cover a broader array of activities, thus lessening the possibilities for error to be introduced elsewhere in the process. As such, our review concludes that the Bureau's program, as reflected in these efforts, was both comprehensive and extensive, and, as designed, met the goals of protecting adequately against failures in the conduct of the discrete operations. Similarly, the written materials, both manuals and training guides, as reviewed, seem fully adequate. However, we did note some comments that greater emphasis on clarity in presentation might have prevented possible misunderstandings and errors in applying the procedures. To that end, we would suggest the need for more extensive testing of materials, as well as of training approaches, to ensure that the materials are clear and unambiguous, and that the range of possible situations to be encountered is anticipated, and provided to the users.

6. THE DEMING PHILOSOPHY

The Bureau first sought to adopt the Deming philosophy in its 1990 Census QA programs, and subsequently into Census 2000. As part of its evaluation of the QA program, Westat was asked to examine the appropriateness and applicability of Deming's management philosophy to the conduct of a decennial census, a very large-scale single event conducted in a very short time period. Any attempt to evaluate Deming's management philosophy in relation to the conduct of a decennial census must first note those aspects of a decennial Census that make it very unique and which lead some to question the relevance of Deming's approach, and, next, highlight the specific management points that are directly germane to such a statistical undertaking.

6.1 Unique Aspects of the Census

There can be little argument that the Census is unquestionably a unique "production" process, a massive challenge that occurs only once every ten years. Over 650,000 field staff are hired, trained, and supervised out of more than 500 local field offices while conducting their tasks. Most of those hired complete their employment in less than six months. Following an immense logistical undertaking of mailing (and, in some cases, delivering) the appropriate census forms to more than 100 million households, some 42 million households are contacted personally, in some cases up to five times, in an effort to complete the data collection process. The job, then, is huge, both in the number of workers involved and in the geographic area over which they are spread. The entire enumeration effort must be completed in approximately eight months, from hiring to exiting. In addition, the process occurs so infrequently that few of the staff involved in the previous Census participate a decade later. Simply put, one finds very little, if any, of the usual "memory" involved in repetitive operations conducted, over time, by essentially the same people.

The Census also is unique in how it attempts to manage its human resources, i.e., how it supervises this vast army and communicates its needs and methods. A majority of the staff has never worked for the organization before and never will again after this enormous, short, and intense effort. Therefore, many of the opportunities for "on-the-job" training are not present, and staff has little time to develop loyalties to the organization, its managers, or to each other, or to learn how to work with their co-workers.

Because of the “one of a kind” nature of the decennial Census, it is reasonable to ask if management practices that have demonstrated effectiveness in business, industry, and education, or in the private sector and government would apply.

6.2 Summarizing Deming

Although Deming’s management philosophy contains 14 points (see Appendix B), we focus on those selected few that, in our judgment, are most appropriate and relevant to a government statistical agency and, especially, to Census 2000. We begin with the more general that apply, namely to “break down barriers between groups,” to “provide leadership,” and, to “drive out fear.”

“Barriers between groups”, if not dealt with appropriately, directly impacts quality. Although far from unique to the Bureau, as we note in some detail later, such barriers appear to have been present and to have played a large role in the difficulties encountered in carrying out Census QA responsibilities, both in Census 2000 and in earlier censuses. Dillman’s paper (1996) noted that barriers existed between different operating units at the Census Bureau, and described the typical government agency as having a “stove pipe” organizational structure designed to channel communications inefficiently up the line, across and down a different pipe. Not only did Dillman find this inefficient, he also observed that it could result in garbled messages caused by too many interactions between the source and the needed destination. We observed situations where this system actually prevented staff from trying to communicate with staff in other departments. Dillman also noted (as did we) a tension between the QA staff and Operations, a tension that has been observed in other statistical organizations as well. Reducing these barriers should contribute significantly to the quality of the Census results, as well as to the smooth functioning of the entire process.

“Leadership”, another of Deming’s points, is that intangible quality which lifts people and staff above the ordinary. In his book, *The New Economics* (1993), Deming devoted an entire chapter to the subject of Leadership, noting that the job of a leader is to accomplish transformation of the organization. Deming realized that only the most senior staff are in a position to exercise the leadership needed to constantly improve quality. Conversely, leaders who continuously ask about production, schedule, and costs, and rarely ask about processes, quality, or quality improvement, quickly demonstrate to staff which priorities are truly important to them. Deming provided an example of a manager who, in his view, demonstrated organizational leadership for improved quality—Morris H. Hansen, who spent over 30 years at the Census Bureau and served as Assistant Director for Statistical Standards. Hansen played a major role in assuring the quality of several Censuses (1940, 1950, and 1960), and left a strong legacy of competence and quality. Deming’s choice of Hansen, whether serendipitous or not,

demonstrates clearly that Leadership, if applied effectively, can ensure that improved quality is seen and accepted as a worthy and attainable goal of a statistical institution. The successful completion of Census 2000 is testimony to the importance of such leadership. At the same time, some of the difficulties in implementing the QA program, which are reflected in this Report, also illustrate a need for greater leadership.

“Fear” is all too often a forgotten element in the conduct of an operation, but it also is highly relevant in any discussion of a decennial census. It is only logical to acknowledge that fear or uncertainty must be present among the many hundreds of thousands of new employees, whatever their level or degree of responsibility, hired for a relatively brief period, to take part in an unfamiliar host of activities previously unknown to virtually all of them. The challenge for the Bureau is to quickly build a working environment in which fear, real or latent, is minimized, thus allowing the new staff to focus on the task at hand and ahead, and to be open to the need for quality, adherence to instruction, and amenable and responsive to supervision and review. Fear can prevent communication, inhibit staff from questioning dubious or unclear instructions, and dissuade staff from identifying problems. Although we were unable to speak with temporary staff and confirm this concern, we believe it is of sufficient importance to the Bureau’s goal of assuring quality to be noted and addressed.

One of Deming’s most important recommendations, which is directly relevant to a census operation, is “to build quality in”, by which Deming meant to use statistical methods to plan and monitor not just the product but the process as well. This is the major distinction between QA and QC. When applied to a decennial operation, the goal is to distinguish the unique or special variation or cause from the repetitive or common cause, and to change or modify the process to minimize or remove the common cause. Thus, if large numbers of interviewers are found to be filling out a listing form incorrectly, the problem undoubtedly is not with the interviewers, but rather with the instructions or the forms; the cause should be isolated and the process corrected, rather than blaming the problem on poor performance by interviewers. Given that some two years after the completion of data collection, neither much of the QA data collected during Census 2000, nor data on the effectiveness of the QA programs, were available to us, highlights a significant, missed opportunity.

The timing situation in a decennial census certainly magnifies the need to understand quickly what the cause of the problem is, and whether it is common or special. In fact, there is a general misconception that there is little that can be done to change the basic procedures or modify on-going operations once the process is underway. Morganstein and Hansen (1990), however, suggested a different viewpoint, namely, that the only hope a Census manager has of influencing quality during this short period is to monitor processes and to use statistical thinking to determine, in real-time, if a special cause

has occurred that warrants a prompt management reaction. Thus in the previous example, the listing form should be corrected during data collection, not after errors have been made on all the forms. We are fully aware that modifying an operation in midstream is difficult, but the alternative, ignoring the problem, is both self deluding and defeating. In the well known Hialeah situation, not only were the offending staff not identified while they were curbstoning data, they subsequently were assigned to other areas as “problem solvers”, because they were incorrectly thought to be accurate and efficient!

Awareness and understanding of problems and their cause is possible only if a system exists for acquiring accurate and timely information on the potential sources of special cause variation. Designing, testing, and implementing the required real-time reporting systems is an integral component of “building in quality”. We believe that this an area in which the Bureau was seriously lacking in its ability to identify and quantify specific types of problems in a timely manner.

Another of Deming’s points, which apply to a decennial census, is “to cease dependence on inspection.” It is apparent that it is not possible to inspect all of the massive elements of information collected throughout the decennial process. Thus, the Bureau wisely utilized statistical approaches to examine a sample of an interviewer’s work to gauge the quality of the work performed in the entire assignment area and to set thresholds for accepting or rejecting the entire assignment. These acceptance-sampling thresholds should be reviewed throughout the data collection process and, based on the findings of the ongoing QC activities, revised accordingly. The Bureau is required by law to produce the population count and its associated characteristics, and it must ensure the highest degree of completeness and accuracy possible. Clearly the Bureau’s legal mandate can not be accomplished by sorting, identifying and eliminating errors after the fact. A major part of the Bureau’s strategy must be to insure quality results through adequate planning, comprehensive training, and careful monitoring.

These, then, are the aspects of Deming’s Management Philosophy, which, in our judgement, are directly applicable to the conduct of a decennial Census. In subsequent Sections, we will address how effectively these precepts and points are perceived to have been employed by the Bureau, and how the programs to accomplish the “Deming Philosophy” are perceived to have been planned, developed, implemented, and reviewed. Finally, we will attempt to assess the Bureau’s strengths and weaknesses in the QA area and to provide some recommendation for consideration in planning the 2010 Census.

7. A SUMMARY OF VIEWS ON QUALITY ASSURANCE

This section presents facts, views, reactions, and opinions about QA in Census 2000, expressed by a range of staff involved with or knowledgeable about aspects of the Census 2000 QA program.⁶ This information was obtained in a series of interviews that were conducted with census staff members, either alone or in groups, and through our review of documents and other materials concerning Census 2000. Our assessment of the QA process in Census 2000 appears in Section 9.

As we noted earlier, although our initial focus was on the eight major field operations described in Section 5, the discussions became more broadly focused on the overall field QA program, and especially as carried out during the decennial collection period. In addition, we found that relatively little data concerning the usefulness and success of the individual QA programs were currently available, although such information is in process of being compiled. Accordingly, for the most part, the comments shown below are directed to recollections of the effectiveness of the QA programs during data collection, as well as an overall, broad overview of QA in the data collection phase of Census 2000. It should be emphasized that some of these comments reflect “perception,” whereas factual reality, indeed, may be somewhat different. For these reasons, we want to emphasize that, for obvious reasons, we were limited in the number of persons with whom we could discuss these issues; nonetheless, we have been careful to ensure that the views reflected below are generally shared (even though we did not attempt to count responses or provide distributions) or are the considered conclusions of a person in a position of significant authority or responsibility and, thus, reflect direct, accumulated knowledge and important opinion. As such, they provide important insights into the attitudes towards, if not the reality of QA, and should not be dismissed summarily. Our discussions evoked both positive and negative comments, sometimes on the same issue or program, and even from the same discussant. We also would note our surprise at the extent, and relative consistency, of the negative reactions and comments, or the dichotomy of view, concerning the QA effort, forthcoming from those with whom we spoke. Although positive comments about the QA efforts were expressed, and are reflected in this section, we believe that most participants were identifying areas requiring improvement and that their comments addressed that view.

- The Census 2000 QA program was perceived to have two objectives, namely:
 1. Content quality—Ensure the quality of the Census; and
 2. Face validity—To allow the Bureau to demonstrate to Congress and others that QA programs existed as part of the Census;

⁶ See Appendix A for a listing of Census Bureau staff contributing the information and comments for the section.

- There was general agreement among those interviewed that the Bureau allowed much of its institutional memory of QA activities during the 1990 Census to disperse following the completion of the 1990 Census and prior to the initiation of planning for Census 2000. Although not unexpected and similar to actions taken at the conclusion of previous censuses, this action is seen as a serious flaw in Census 2000, partially because Census 1990, correctly or otherwise, was seriously criticized by the Congress, in the Press, and by selected users. The resulting intensified efforts to start the Census 2000 planning at an early point in the decade found the Bureau short of “institutional” knowledge in the QA area. Most of the staff who had worked on the 1990 QA effort was dispersed prior to the initiation of planning for Census 2000, and a centralized QA leadership was not established at an early point to plan QA for Census 2000. In fact, early planning for Census 2000 decentralized the QA function to “task groups” chaired by the Field staff. All of these developments, taken together, are perceived as resulting in a significant loss of QA staff expertise, staff continuity, overall staff size and, most important, leadership at an important point in the planning for Census 2000;

Further, the decentralization itself is viewed as having contributed to a perceived lack of an established, overall Bureau philosophy regarding QA until late in the decade. In fact, to our knowledge, the only QA Seminar to inform staff of the Bureau’s QA philosophy was not held until June 1999. For the most part, during the decade, each working group developed the QA program for its assigned function in a somewhat isolated situation, without any apparent centralized coordination or review.

- Preparation for Census 2000 does not seem to have built on or benefited to any great extent from the results of the evaluation of Census 1990. To a great extent, this is ascribed to the fact that most of the 1990 evaluation studies merely reported on the errors found, with little attention to implementation problems;
- Limited developmental QA research took place during the intercensal decade, with the emphasis on “cost savings,” through improved operations and changes in project management;
- The initial planning for Census 2000 did not include an overall QA plan, nor were separate budgets provided for the activity; rather, each program was assumed to include funding for such QA as might be required. Thus, early funding was available only on an ad-hoc basis;
- A measure of the importance of QA, as perceived by virtually all of those charged with its implementation, can be summarized as follows: The QA initiatives were the first things cancelled or curtailed, and the last things allowed in;
- Although a small, centralized QA group, consisting of only five persons, was reestablished in 1998, key responsibility for approval of the approach and subsequent implementation of QA remained with the operational entities, which limited the role of the QA staff;

- The change in census operation ordered by the Supreme Court in January 1999, which forced the Bureau to shift from the planned “Integrated Census Method” approach to the standard Census approach, was particularly difficult for the QA program. With the need to revise so many aspects of the overall decennial operations in a relatively short time period, few resources and little time were available for changing the QA approach;
- Although some QA operations were included in the testing between the 1990 and 2000 censuses, the QA operations were not the focus of the test evaluations; further documentation of the QA testing or QA results from the tests is lacking. In addition, no analysis of the Census 2000 NRFU Dress Rehearsal Reinterview program was undertaken. Finally, due to time constraints, revisions or modifications to the QA programs were implemented without any further testing;
- Responsibility for developing and designing QA programs and measuring instruments, for the most part, was delegated to the QA Branch in DSSD; implementation, of necessity, was the responsibility of the Field. During the data collection effort, as a result, problems were dealt with at the local office level and rarely reached the QA staff or rose to a higher level. Involvement of the Executive Staff appears to have been minimal;
- The QA staff was perceived by Field headquarters as impractical and not aware of the realities of Field requirements; in the words of one senior official, “with their heads in the clouds.” More generally, they were viewed as a team of scientific experts demanding real time information for immediate reaction, not recognizing the reality of a Census, and totally insensitive to the programmatic needs and requirements of data collection. To paraphrase, “the QA staff must realize that at the local office level, staff is focused on production; in order for QA to be perceived as important, it must be ‘built into the process,’ which is very difficult and too easy to circumvent”;

Field, in turn, was perceived by the QA staff as unwilling to recognize any role for the QA staff, especially once data collection was underway.

- The role of the QA staff in implementing QA programs in the Field was quite limited. Implementation rested completely and solely with Field. Field had “final” authority on all aspects of QA, including program development, decisions on training materials, control of travel to observe field activities or operations, development of OCS2000, real-time access to the information from OCS2000, involvement with RCC and/or LCO leadership, and to problems;
- The Field staff lacked independent QA management at all levels. Specifically, neither the RCC nor the LCO had any dedicated QA staff, with the result that QA, effectively, had no independent voice and was subordinated to production;
- Although exposed to the need for QA and the QA programs during their initial training, for the most part LCO managers, not surprisingly, found their greatest challenges elsewhere, and paid little attention to QA concerns. They understood, in theory, the need for “quality,” but “in practice,” “production” took precedence over quality;

- QA consists of “Prevention” (do it right the first time), and “Rectification” (provide information quickly and efficiently to catch mistakes and problems). There was general agreement that QA as practiced in Census 2000 had mixed results—it was a definite contributor in preventing errors on the part of interviewers, but it was generally too late to affect the rectification goal. The time allowed for a given Field operation, for the most part, was seen as far too limited to allow for QA results to affect the operation, or result in a change in procedure, and, in Hialeah, it failed on both levels. For example, in Hialeah, an office which had significant problems in undertaking NRFU, the NRFU Reinterview was virtually complete by the time questions were raised, and such QA results were far too late to have alerted either the Regional Office or Washington.

It was pointed out, however, that the QA program could have identified the problems in Hialeah if there had been adequate management of the QA operations at the LCO. In addition, real-time data available to Field management or other levels of review were wholly inadequate. For example, MIS reports failed to provide any information on cases passed or cases failed. Further, progress data for a date some 2 weeks after Random Reinterview should have begun, showed some 50 percent of the NRFU workload reported as completed, as compared with only about 38 percent for the NRFU reinterview workload. In so far as the goal of the Random Reinterview was to identify falsification and other problems at a very early stage, these rates should have shown a reverse relationship.⁷

- However, notwithstanding Hialeah, which Field prefers to describe as an “outlier,” the prevailing Field view is that virtually all local offices used the results of QA in “real time,” and that the collected data benefited from the QA effort. Put in context, QA programs designed to identify interviewers who required additional training or who failed to understand their tasks generally were successful in meeting their objectives in most of the stages of the Census prior to NRFU, but much less so in NRFU. Further, many of the problems occurred in attempting to determine the extent of the problems, such as through reinterview. NRFU QA resources were inadequate, partially because the Bureau underestimated the extent of followup. This is borne out by the million-case backlog in NRFU reinterview by the third week of scheduled activity, as reported in an OCS2000 Report. This large backlog also suggests the failure of the random reinterview, which had the primary goal of providing “early” protection against falsification and poor performance, resulting from not having been implemented as designed. Essentially, it appears that the QA program was just too ambitious, especially given that management compressed the time schedule for NRFU in order to meet the overall decennial time schedule. Thus, the QA/NRFU Reinterview became a casualty of the time schedule;
- Because of the foregoing, the QA component of NRFU, the largest and most complex of the Field QA efforts, is seen by many, as well as by debriefing comments and the Bureau’s own survey of selected field staff, as generally having been done poorly by most local offices. To repeat, some view this result as a failure to provide adequate resources. Nonetheless, the result is the same. Field, however, does not agree fully with that assessment, believing that, on the whole, NRFU accomplished its objectives. Again, data are lacking at this time, either to confirm or deny either assertion, although an effort is underway currently to compile the relevant information.

⁷ Based on information developed by field from the OCS2000 system.

Whatever the outcome of these reviews, however, the Bureau itself noted that “reinterview program expectations were not clearly communicated to all field staff and caused confusion.”⁸

However, one piece of information on the success of the NRFU QA effort is available from a survey conducted by the Census Bureau subsequent to the completion of NRFU, among NRFU Office Operation Supervisors.⁹ The comments of these supervisors raise some serious questions. As noted in the Report: “a majority of the Office Operation Supervisors (OOSs) reported that they were not given sufficient time to perform the functions. Their Assistant Manager for Field Operations (AMFO) simply gave them self-study training guides and a few days to set up the operation. To make matters worse, many of the OOSs reported that the reinterview material arrived late or had missing pieces. Some of the manuals and job aids did not contain sufficient information so they had to fill in the missing pieces through trial-and-error.”

They also reported that by the time the program was set up, the NRFU enumerators had completed a significant portion of their respective NRFU workload, rendering the Random Reinterview much less useful than planned. Confirming some of the comment noted earlier, they also noted that by the time the FOSs had returned their reports designating the enumerators whose work was to be reinterviewed, the enumerators either had completed their entire workload or had quit. Finally, they observed that when the reinterview workload conflicted with the NRFU production deadline, reinterview often went by the wayside. As an aside, it should be noted that the decision on designating enumerators for reinterview was not the responsibility of the FOS, which illustrates how the independence of QA was compromised due to a lack of QA management at the LCO level.

These reactions also apply equally to the U/E and L/E reinterview operations.

According to the OOS survey, the Reinterview operations “did not work smoothly as we would have liked. Many LCOs started this operation late or did not execute it properly. Some LCOs did not perform Reinterview;”

- We also would note that the Evaluation Requirements Section in the Program Master Plan for NRFU is significantly lacking in questions concerning the effectiveness of the QA program. For example, it appears that no information was collected concerning the effectiveness of the QA program in identifying “poor” or “inadequate” interviewers or its effect on data quality, including the extent of interviewer or clerk replacement resulting from QA failure;
- Census 2000 was the first Census to “overemploy”. The Bureau hired at almost a 200 percent level for enumerators, thus placing excessive pressure on the hiring and training capacity and ability and, subsequently, on the capability to QA/reinterview the staff. The overly large staff of enumerators overwhelmed the ability of the supervisory staff to train, observe, and monitor performance;

⁸ Census Bureau. (2002). *Assessment Report for Nonresponse Followup*. Final Report. Census 2000 Information Memorandum #127, page 7.

⁹ U.S. Census Bureau, *Lessons Learned from the Census 2000 Nonresponse Followup Reinterview Operation*, March 2001-May 2001. Unpublished.

- In the words of one senior Bureau manager, however, “QA was successful, and worked better than in past Censuses”, for the following reasons:
 - It provided Crew Leaders with edit information;
 - It gave Crew Leaders management information on such items as incompletes, and inconsistencies between the population count and questionnaire pages; and
 - It was a visible program.
- Stated in somewhat different terms, “QA was successful, not 100 percent, but at least 75 percent, and it was at least 10 times better than in 1990.” “Census 2000 provided a separate staff for reinterviews, whereas Census 1990 required Crew Leaders to conduct this program.” “In the view of top management, the QA 2000 program wasn’t very bad; “on a ten point scale, we aimed mostly for eight and probably got only to six”;
- Nonetheless, a dichotomy of views exists. Simply put, QA staff view the cup as half-empty, Field staff see it as half full, with the Bureau getting better and better at looking at and finding out where it fell short; this view seems to be shared by top management;

Management also believes that if technology and automation can replace paper, the QA program for 2010 will be simplified, easier to implement, can start earlier, and can get in on the ground floor; and

- Despite the many QA problems encountered in Census 2000, many share the view that the Bureau has gone a long way towards adopting the QA philosophy of involving the entire organization. Its effect can be seen—even if not fully recognized or acknowledged—in effective QA programs for the printing of forms and the preparation of maps, in higher supervisory/staff ratios, better testing, hiring and evaluation, and more resources made available for the planning effort. However, problems continue to exist at the implementation level. Also noted is the fact that Census 2000 had some form of QA process in place for most operations, unlike Census 1990, which had a much more limited QA program.

These comments by staff who played diverse roles in Census 2000 should provide valuable input to and guidelines for the planning for Census 2010.

8. QUALITY ASSURANCE IN CENSUS ACTIVITIES IN OTHER COUNTRIES

As part of our review of QA activities in Census 2000, we also were asked to examine the approaches used by several other countries collecting census data. After some investigation, we chose the United Kingdom, Canada, and Australia, each of which had completed a population census in recent years, and also had an approach to census taking somewhat similar to that of the U.S. In addition to obtaining information on the kinds of quality assurance programs established, we attempted to develop some indication of how well they were carried out.

On a more basic level, we were curious about the philosophy behind the approach used by each of these countries and how this philosophy led to the selection of the specific programs. We also attempted to explore the reasons for differences between the programs used in the other countries, as compared to those used in the U.S. In effect, we attempted, in summary fashion, to obtain some information about the respective QA programs and efforts as we did for Census 2000. Finally, the key objective of this effort was to evaluate these programs from the point of view of their suitability for Census 2010. The information which follows is based both on discussions with staff of the respective organizations and on review of published materials concerning their Census activities.

8.1 Office for National Statistics, United Kingdom

The most recent census of population in the United Kingdom took place in April 2001. It was the twentieth census to be carried out in Great Britain and the eighteenth to be carried out in Northern Ireland. Our investigation was limited to the census of England and Wales, which was carried out by the UK Office for National Statistics (ONS). England and Wales were divided into 103 Census Areas, each of which was managed by a Census Area Manager (CAM). In turn, there were some 2,000 Census District Managers (CDMs), some 6,000 Census Team Leaders (CTLs), and about 62,500 enumerators to carry out the enumeration of the 22 million households.

Prior to the Census, a list of all addresses was prepared, using Post Office lists and other available sources. Maps noting the addresses within Enumeration Districts and Enumeration Record Books listing the addresses were prepared and provided Enumerators to assist them in locating households. Each map was assumed to cover a single Enumerator's assignment, approximately 400 addresses (200 in inner cities). In early April, before Census Day (April 29, 2001), Enumerators, using the list developed earlier, identified, verified, and visited every address to deliver a Census form and mail-back envelope. As necessary, missed addresses were added and those no longer in existence were deleted.

Residents at each address were asked to complete the Census form and, for the first time, to return it by mail. Addresses for which forms were not returned were revisited by Enumerators, who collected the form directly, if available, or asked the householder to mail it back as quickly as possible. Enumerators also visited addresses to obtain missing information, as required.

The mail-back response rate far exceeded expectations; in all, some 88 percent of the households returned their forms by mail, well above the expected 70 percent. This larger than expected mail return caused some disruption to the enumeration, as the unexpected flow seriously overwhelmed the postal service.

The goal of Census 2001 was to maintain the high level of coverage achieved for the majority of the population in the previous census (1991). To this end, resources were concentrated on improving the coverage of the population in the groups that proved hard to enumerate; QA efforts, not surprisingly, were focused to a large extent on this objective. For example, an extensive community liaison program was established to reach out to residents of difficult areas. In the data collection phase, this objective was accomplished through a strategy of instilling a concept of “Team Work,” which attempted to motivate all staff with an understanding of the need for and commitment to quality.

The following QA programs were included in the data collection phase:

- The CDMs conducted a “Census District check” of the addresses prior to their use by Enumerators, both to familiarize themselves with the area, and to catch any obvious problems of omissions, boundaries, mapping, or resource allocation. This activity also served as a means of alerting senior officials prior to the onset of data collection of any problems with this phase of the process. All of the local governments in England and Wales also were asked to identify all major housing changes, whether demolition or development;
- Significant QA efforts were devoted to ensuring that the recruitment process successfully identified the best candidates. Extensive training, consisting of self-study and testing, was used to provide Enumerators with an understanding of the task and its elements. Finally, CTLs observed the Enumerators early in their delivery of the forms and provided such feedback and/or additional training as seemed necessary. The Field checks were used to emphasize the importance of finding all addresses and of staying within Enumeration District boundaries, as well as setting the framework of what was expected of the Enumerators on the ground. The CTLs also reviewed the Enumerators Record Books to ensure the proper completion of required entries, and prescribed action was taken, as required, if the review showed unacceptable results;

- With the return of forms through the Post to the local Census offices, CDMs and CTLs were responsible for dealing with the receipt and checking of the forms; supervising the followup visits by enumerators to addresses from which forms were not received; and carrying out a ‘mop-up’ of nonresponse. The check-in consisted of a “quick flip through,” with failures returned for a field followup. There was no systematic QA program or effort, such as a reinterview, to ensure that followup was done properly. The emphasis at this point was on ensuring that a form was obtained from every household, and that every household was accounted for; and
- CDMs conducted a completion check on all forms. Incomplete forms were identified based on a very limited set of key questions (e.g., sex, date of birth, and marital status).

Although QA was part of the regular training, the importance of the QA checks and the reasons for them were emphasized to all Field staff. Enumerators were told, with emphasis and up front, that their work would be checked.

The Field Management Information System (FMIS), which was established to monitor progress and provide close feedback between the Census Office and the field staff, was not designed to provide any quality measures, nor was it used to alert supervisory levels about QA problems. CTLs or CDMs generally dealt with such issues. Further, the MIS system occasionally was unable to provide information on a “real time” basis because of system failures.

Given its emphasis on coverage, rather than content, the “final” quality of Census 2001 will be evaluated by results yet to be obtained from a post-enumeration study of some 300,000 households, conducted some 4 weeks after the completion of the Census. Results of the Post-Enumeration Survey (PES) survey will be used to adjust numbers of both people and households found in the Census.

Turning more directly to the QA process, the QA strategy was established originally by the Census Program Board, chaired by the Director of the Census. Subsequent decision making was decentralized to lower levels of the organization, although significant issues did move up the chain of command quite rapidly. The specific QA programs were developed by the individual operating groups themselves, but a “data quality manager,” reporting directly to the Deputy Director of the Census, had an oversight role and advised on, if not approved, the specific programs. This position was filled some five years in advance of the Census.

In the Census of 1991, some friction existed between the field operations planning staff and those charged with carrying out the QA effort; for that reason, a great deal of focused effort went into the planning for Census 2001 to establish a common goal among the participants and a commitment to

cooperate. Information technology, statistics, and field operations staff were placed in a single location. The staff selected was a mixture of both experienced and new people with open minds, with an emphasis on listening and the ability to compromise.

Although the QA effort did not have separate funding, it is estimated that between 5 and 10 percent of the 207 million pound cost of Census 2001 was allocated to QA.

Reflecting on the QA effort in Census 2001:

- In hindsight, the QA effort is seen as a bit of both a QA program and a QC program. The QA component is reflected in the coordinated planning of the programs to achieve the desired level of quality; however, the individual programs are best seen as a QC effort;
- QC training was incorporated into the regular training program. The importance of quality checks was explained, and Enumerators were advised that their work would be checked;
- There is no information currently available to determine whether the program achieved the “desired” level of quality. The results of the Post-Enumeration study will be used as the “yardstick” for such a determination;
- Data were not collected on the number of Enumerators who were terminated, retrained, or otherwise affected by the QA efforts, nor is there a count of assignments that required additional work because of QA concerns;
- The QA program, however, is viewed as having been successful overall, especially given the expected accomplished household coverage rate of 98 percent. Similarly, the program is given credit for a perceived reduction in differential nonresponse;
- The mail-back of the forms and the Census Coverage Survey (CCS) were improvements over procedures used in past censuses, and both developments benefited from the Community Liaison Program. The CCS also utilized a web-based information system that was successful in tracking the progress of the field effort during the survey;
- The FMIS is seen as a failure, which is ascribed to its being considered as a nuisance and, thus, just ignored. Attempting to obtain real-time data for use in managing the Census was said to be “extremely frustrating” and, in fact, was not accomplished;
- The time allowed to conduct the Census, eight weeks in all, may be too short, using a mail back methodology;

- The advent of more sophisticated, yet flexible technology may permit more to be done in a shorter time frame in future censuses. For example, technology may allow for information to be stored in real time and retrieved almost instantaneously for review and followup;
- A formal debriefing process was implemented, in order to retain the experiences gained in Census 2001. All CAMs were directed to provide views and opinion on and reactions to a wide variety of procedures and programs, including QA. In turn, they had debriefed their own teams, so the cumulated information is expected to be quite comprehensive. The results are being placed in a data bank for future reference. A wide variety of ad-hoc memoranda on census issues and problems also are being incorporated into the database; and
- The Census 2001 QA approach is seen as “Total Quality Assurance,” and the programs were set up accordingly to accomplish this objective. As noted earlier, however, the focus was on coverage and differential undercoverage, with apparently little attention focused on content or on within household counts.

8.2 Statistics Canada

The 2001 Census was the 19th in a series, dating to the first national census of Canada in 1871. In 1956, Canada began conducting a census twice during a decade—in the Years ending in “1” and “6.” Census 2001 involved some 12.5 million households, containing about 30 million persons. Some 80 percent of the households were asked to complete a short questionnaire, containing seven questions; the remaining 20 percent received a long questionnaire, containing 59 questions, in all. The forms were distributed to households in 45,000 Enumeration Areas (EAs) between May 1 and May 12, by some 34,000 Census Representatives (CRs), to be completed as of May 15, which was Census day, and returned by mail. As part of the distribution process, CRs created a list of all private dwellings, known as “the visitation record (VR).” A mail return rate of 85 percent was achieved. In areas without good addresses, CRs collected the information directly from the households.

The CRs were hired, trained, and supervised by approximately 2,800 Census Commissioners (CCs) who, in turn, were hired, trained, and supervised by CDMs. The field structure also included regional Census Managers (CMs), CAMs, and QC Technicians, who had responsibility for ensuring that the QA activities were carried out correctly in local offices.

In advance of the Census, an Address Register, containing a listing of all households in municipalities of 50,000 or more persons, was prepared. The listing covered only about 63 percent of the total dwellings in Canada. The listing consisted of addresses recorded in the Census of 1996, updated by adding potential new dwellings identified through administrative sources. There was no local municipality review of the final Address Register listings. The Address Register, however, was not used

for delivery of the Census questionnaires. Rather, as noted, Census 2001 methodology called for the enumerators to list all addresses at the time of questionnaire delivery. The Address Register, in turn, was used as a coverage improvement tool through a reconciliation process, first introduced in the Census of 1991. After enumerators had completed the listings for the assigned enumeration areas, they were provided with the Address Register Booklets, which contained the list of addresses for the same areas. The enumerators then compared the two listings and verified any addresses that had been missed during the listing operation. Questionnaires were completed for missed households.

Followup was of two kinds, the Failed Edit Followup and the NRFU. At the time the form was returned by mail, the CR conducted an edit consisting of adding up the number of nonresponse questions. If the number exceeded a pre-established level, the questionnaire was considered to have failed the edit and a telephone followup was required by the CR to complete the form. Mail nonresponse followup, which began several days after Census Day, also utilized telephone followup, if feasible. Four separate attempts by phone were required, after which three visits to the dwelling were required. Dwellings without phones were visited directly.

A factor which could have affected quality was staff turnover, which resulted from the relatively low rate of pay, and which reached close to 45 percent for CRs. However, Statistics Canada, fortunately, had a relatively large reserve pool of qualified applicants readily available, and was able to replace CRs as needed. This development, however, did require shifting rather significant resources to the training of replacements.

The QA program for Census 2001 was based on the following decisions established during early Census planning:

- Everyone would be responsible for quality. Each level of activity would be expected to produce some sort of a Quality Report to document all activities and all decisions;
- In order to optimize the effectiveness of the work force, a centralized hiring system would be developed and utilized;
- The experience from past censuses would be used to develop better training methods, including the use of more electronic media;
- Every document would be controlled, and “failure” thresholds would be established for implementing followup;
- CRs would be given the responsibility for safeguarding against bad quality data. They would serve as the “first line” of defense;

- The MIS would permit the pinpointing of areas where nonresponse was high, allowing management to move quickly and assign resources to alleviate the problem; and
- An audit of QC would be required for Census 2001, in order to determine if the specified QA procedures were carried out correctly and appropriately.

The activities designed to implement the foregoing decisions and to accomplish QA in Census 2001 included:

- Following the drop-off operation, CCs conducted a one-day training, during which the CC verified that all CRs both understood and applied the procedures correctly. The CC also used the opportunity to allow CRs to share their experiences with other members of the group;
- Every listing in the VR was required to have an acceptable disposition entry, including the presence of a completed census form or proper documentation for a vacant dwelling, a temporary resident dwelling, or one containing foreign residents;
- A nonresponse tolerance level was established. If the percentage of households with no questionnaires exceeded 1.8 percent, the EA was assigned to a “clean-up” operation, consisting of additional efforts to contact the nonresponse households and complete the census forms;
- Checking followup attempts. When the percent of incomplete forms exceeded five percent, a supervisory check was undertaken to ensure that CRs tried to contact all of the households lacking resolution. If at least one address with an incomplete census form was found to have no indication of a field followup attempt, the entire EA was failed and reassigned;
- A sample of documents was checked to determine that the controls were completed properly. If too many mistakes were observed, the EA was redone and every document recontrolled and followed up appropriately;
- The CCs had a comprehensive list of items to verify at the time the CR returned an assignment (EA) as completed. After approval by the CC, the EA was then reviewed by a QA Technician. If approved, the EA was then sent on for processing; if rejected, the EA was returned for further followup. Although each Region had a degree of autonomy in how to implement the different requirements, the QA standards were centrally established and implemented accordingly. EA reject rates varied considerably across Regions (from 6 to 25 percent), but these rates did not reflect the final quality. Rather, some Regions conducted their reviews and rejected the EAs as soon as they were received, without any attempts to correct errors at that point; other Regions attempted to correct EA problems before considering them for rejection;
- An MIS system was established and utilized to identify problems in the data collection process. Unfortunately, the MIS was unable to meet the more important needs for item detail, such as the number of errors by question; rather, the MIS was limited to gross process flow results, such as the number of EAs ready to be shipped and the number failed. Even cumulative data were not retained on the MIS.

The MIS problem is reflective of the tensions between different groups within the organization; in this case, the production side, as opposed to the QA group. The production group saw the MIS as a “Field” tool, designed for and restricted to their unique needs and requirements. Those responsible for QA, the Methodologists, looked to the MIS as a vehicle to provide on-going, real-time information at a specific detailed level.

- Procedures were established for each of the centralized Field Collection Units (FCU), which ensured that VRs were complete, all units were accounted for, and questionnaires were acceptable for processing.

In order to ensure that the “lessons” of the current Census were “learned” for future Censuses, each level of supervision was responsible for preparing Quality Reports which documented their activities and decisions. In addition, studies were implemented as part of the Census to evaluate the impact of the control and followup operations on the collection phase. Studies also were incorporated to evaluate the effects of different collection methods on quality of the results. Subsequent to the Census, debriefing sessions were held at each level, and recommendations were compiled concerning all phases of the operation, including QA. Responsibility for these efforts fell under the Collection Methodology Task, established with a mandate of providing evaluations of big changes in procedures, with a goal of looking ahead to future collection methods.

As a final point, the ability to maintain “institutional memory” was especially emphasized as a key factor in improving Census quality from one Census to the next. Conducting a Census every 5 years has allowed for substantial continuity at the managerial level in the Census organization. Specifically, a vast “census culture” can be maintained, although there is full awareness of the need to add staff selectively to prevent “bad inertia.”

A Reverse Record Check Study is used to evaluate the coverage of the Census. Effectively, a sample of names is selected from the previous Census and supplemented by births and immigrants arriving during the intercensal period, and an effort is made to locate that sample of individuals in the current Census.

In retrospect, the VR check is seen as one of the particular strengths of the recent Census, in that it required the disposition of each listed address. The one-day training session held immediately after the distribution of forms also was seen as especially effective in dealing with the misunderstandings or questions posed by CRs and, also, in motivating them to accomplish their task. The key weakness was that EAs were found to be too big, that is, to contain too many dwelling units and, thus, the time required to complete the EA delayed and dragged out subsequent operations. This problem also required the QA

review to be cut short in many cases, in order to meet the established time schedule; instead the EAs were sent directly to “clean-up” for data repair, which maintained quality, but at increased cost.

With Census 2001 results still to be released, it is far too early to attempt to assess fully the effectiveness of the different QA program or to suggest possible changes for future consideration. One plan already underway, however, calls for developing a centralized system for maintaining data for the Census of 2006. Such a system would contain the entire database, as submitted on a flow basis. As such, all components of the operation could access the common database, as needed, to obtain the types of diverse information required to control and manage their respective responsibilities.

It also is possible, at this relatively early stage in the processing of the Canadian Census of 2001, to note that QA is an important and visible component of the Census during planning, development, and implementation, and consumes significant resources in money, time, and staff. It also is clear that, for the most part, as in the U.S. the early phases of the Census fall in the area of a QA approach, whereas the collection phase is oriented towards QC methods, that is, sample, test, and reject.

It is important to note that the overall QA proposal for Census 2001 was developed within the various teams (collection, coverage, questionnaire design, research and testing, etc.). The broad vision on the coordination for the QA process, however, was driven by the Census Steering Committee, composed of senior management and chaired by the Assistant Chief Statistician. In short, senior management played a significant role in the coordination and approval process of QA planning, whereas execution of the various components was left to the managers of the respective operating teams. However, although the Field hired and supervised the staff which performed the actual QA/QC operations, the direction and reporting on this phase was the responsibility of the Methodology Group, which was independent of Field.

The close and continuing involvement of senior management is seen as a major contributing factor in minimizing, if not eliminating, potential conflicts between the various groups. To borrow the words of a senior manager, “Some tension in a project driven environment can actually be a healthy thing; there is always some push back when one feels there is too much looking over the shoulder. (But) in my view, the consensus buy-in of the senior management steering/oversight committee was a very positive factor in the elimination of a lot of potential conflict in the trenches with the troops. The discussion was very high caliber and tension filled.”

The Canadian QA program is viewed “very much as a combination package”, using both QA and QC, and Statistics Canada would find it difficult to envision a process that was singularly one or the

other. Finally, in their judgement, given what they set out to accomplish, they conclude that their QA/QC program was successful, that there could be significant and substantive improvement but, given the budget and other factors, “it would stand up to scrutiny”.

8.3 Australian Bureau of Statistics

Since 1961, Australia has taken a Census every five years; the latest, Australia’s fourteenth national Census, was taken in 2001. Census Day was August 7th.

A single, “long” form was used to collect the needed information through self-enumeration. A hierarchical structure of temporary staff was used to deliver and to collect census forms from the 9.8 million households and 19.5 million residents. The majority of the workforce, just over 28,000 people, consisted of Census Collectors (about 23,000, in all), who delivered the forms to every household in their collection district prior to census night, and arranged to return and collect the forms or, as required, to complete a form at that time.

The Collection District (CD) was the basic geographic unit of collection, and consisted of a census workload area that could be covered by a single Collector. Group Leaders, each of whom trained and supervised the work of approximately 10 to 12 Census Collectors, were responsible for ensuring accuracy and completeness of coverage within their areas. Where needed (generally in the bigger geographic states), Field Coordinators were used to provide supervisory assistance to the Group Leaders; otherwise the Group Leaders reported directly to the Area Supervisors. In total, more than 32,000 temporary field and collection staff were recruited, trained, and supervised in the delivery and collection of the forms.

The key element in the delivery of the forms to all households was the map prepared for use by Census Collectors. The maps were derived, for the most part, from databases maintained by each State or Territory. There was no subsequent review and revision by local communities. A separate map was prepared for each CD, showing the legal boundary lines for each plot of land in the CD, and served as the collection control mechanism. These maps were used by the Collectors to plan a delivery route, which ensured that the full area was canvassed for dwellings, and that all addresses and living quarters, none of which were noted on the maps in advance, were located and entered into the Collector’s Record Book, which acted as a collection control register.

To ensure that high quality data were obtained from the Census, extensive effort was put into the collection procedures. As a first step, all but the essential administrative responsibilities were removed from both the Census Collector and the Supervisor, leaving them focused solely on the task of taking the Census. Census management believed strongly that this action was critical to improving quality. All QA efforts were designed around a “Philosophy of Quality Management.” The emphasis was placed on each person’s “ownership” of the job, which was reinforced through adequate pay and through training, both classroom and home study, which emphasized strongly that each person was responsible for the quality of his or her own work. The QA processes were put in place to ensure that responsibility was being accepted, rather than as a series of processes to check on quality. The following activities were instituted as part of QA:

- The appropriate supervisor checked the delivery route proposed by the Census Collector in advance of the delivery of forms;
- Collectors were required to scan each form to ensure it had been completed;
- Group Leaders verified that a form existed for each address listed in the Collector’s Record Book. In addition, the number of nonresponses, unoccupied dwellings, and requests for mail-back forms was monitored on an ongoing basis through phone contact and personal inspection of record books and forms. This information also was compared with the supervisor’s personal experience and data collected by other collectors as to “what was possible or reasonable” within the same area. In addition, counts of the expected number of households were developed using the results from the preceding census, along with information from building permit files. If Collectors found a major difference from the expected number, they were required to provide an explanation. Area Supervisors checked totals at a later stage, followed by an automated check at the data processing stage. Independent relisting or recanvassing was not undertaken; and
- A measure of the extent of undercounting was obtained from a PES of households, undertaken shortly after the census.

An evaluation of the data will be carried out to inform users of the data about the quality and to help plan the next census. Another activity aimed at the next census consists of obtaining detailed reports from the different staff levels about their experiences and problems in conducting the Census, undertaking sample surveys of staff, conducting debriefing meetings, and sifting the materials to develop a set of action items and recommendations for the next census. Such action items and recommendations are entered into a database and are then accepted or rejected by project managers for consideration for future action. Accepted action items are automatically recorded on the Census Project Management Framework database for the next Census.

Reflections on the QA effort in Census 2001 follow:

- Explicit efforts to ensure the continuity of staff over the relatively short census cycle resulted in significant benefit to all areas of the Census. For example, the turnover rate among Census Collectors was held to 10 percent. Similarly, demanding, if not insisting on, a collegial relationship between staff in different areas of census responsibility, added significantly to the effectiveness of the QA effort;
- Census top management refused to assume that Data Collectors would fail; rather, they began with the assumption that the Collectors would do a good job and, accordingly, used “trip wire” QA procedures—limited procedures which alerted supervisors and management to major problems in understanding procedures, etc.—rather than imposing a detailed QC approach. They believe this approach has been proven successful;
- In Australia, as in the United Kingdom (UK), coverage results from the PES serve as the final measure of Quality. As regards content, aggregate level data from the census are compared against other sources of such information, such as survey data and administrative record data;
- The Australian public gives a measure of “high acceptance” to the Census, which is reflected in virtually complete cooperation and wholehearted public participation. To illustrate, they note that 100,000 people called to say, “you missed me.” The Census records both very low item nonresponse rates (around 7 percent for income), and misses very few persons, thus resulting in small estimates of undercount;
- Nonetheless, as in other “developed” countries, some problems exist. Specifically, it has become more difficult to contact the young and to obtain entry into buildings that exercise high security. People also are less likely to be home, in general; and
- It also appears that some buildings were missed in the Census. Although some of the missed buildings were caught in the Area Supervisory Review, it is likely that not all missing buildings were identified.

For the Census of 2006, the Australian Bureau of Statistics is planning to use “on-line” reporting systems, which should permit management to access important information on a real-time basis. They also plan to explore the use of the mail-back approach, as well as Internet data collection.

8.4 Summary

Our inquiry into the census taking practices of the UK, Canada, and Australia indicates a keen recognition of the need for and the importance of ensuring “Quality” in the Census. Not surprisingly, it also demonstrates that “Quality” is seen—and interpreted—somewhat differently in different countries. In both the UK and Australia, for example, “Quality,” for the most part, is seen as synonymous with

“coverage”; subject matter content is accorded somewhat less importance. Canada, on the other hand, gives somewhat more attention to content at all levels of the collection effort.

Both the UK and Australia also emphasize “ownership of the job” which, in effect, assumes that staff wants to and will do the right thing, and assigns responsibility on that basis. QA is designed to support and assist that approach. The Australian approach is minimal and hardly intrusive. For example, in describing how it conducts its Census, the Australian Bureau of Statistics devotes a Chapter to “Quality”; which details the importance of form design, collection procedure, field-testing and public awareness, but says nothing of the importance of training the enumerators or of ensuring their adherence to procedure. Similarly, a section on QA deals only with actions taken once the forms are in the census data processing center, but totally ignores the collection phase. The UK, on the other hand, although philosophically in agreement with the Australian Bureau of Statistics, employed more of a QC approach, requiring more interaction and review on the part of the supervisory and review levels.

Neither of these countries, however, utilized any sort of a field verification of the address listing component, nor conducted reinterviews to check either coverage or content. And, in both countries, the PES is seen as the “benchmark” for the Census. In addition, review of content by Enumerators and Supervisors, for the most part, was limited only to ensuring entries to very few of the basic questions which appeared early in the form.

For its part, Canada is both more extensive and more specific in its QA requirements. Although agreeing completely that responsibility for quality is fully shared, Canada requires significant supervisory input and review of the collection process, such as in its 100 percent control of each entry in the VR, and in establishing nonresponse tolerances which, when exceeded, trigger field followup activity. A major distinction in the approach to QA is Canada’s failed-edit review, which examines the extent of nonresponse to the full questionnaire content. Canada also imposes more extensive supervisory oversight of Enumerators and other levels of responsibility, as well as more reporting requirements. To that extent, the QA program in the Canadian Census is similar to that of the U.S.

It is also fair to note that all three of these countries consider themselves to be following the Deming approach. They are extremely sensitive to issues of quality and to the need to institute checks that will ensure their ability to identify aberrations quickly, efficiently, and effectively, and to take those actions necessary actions to improve the system. Yet, they also recognize the time constraints imposed by a census and the limitations imposed on their freedom and ability to make changes. A number of suggestions drawn from the experience of these countries that might be applied to the U.S. process are found at the end of Section 10. In general, however, we would propose that the Bureau arrange with these,

three countries, as well as with any others who engage in somewhat similar decennial activities, to share experiences and past results of QA efforts, and to exchange thoughts and suggestions on future QA efforts in the conduct of censuses. We also would add, “the sooner, the better!”

9. QUALITY ASSURANCE IN CENSUS 2000: STRENGTHS AND WEAKNESSES

The QA Mission Statement for Census 2000 is brief and to the point. Specifically, its goals were:

- To prevent significant performance errors;
- To prevent the clustering of significant performance errors; and
- To promote continuous improvement.

This section summarizes the strengths and weaknesses of the specified QA Field programs as planned and implemented, from the point of view of the QA Mission Statement and a more general view of quality assurance along the lines of Deming's philosophy. However, the section goes somewhat further, in that it also explores the strengths and weaknesses of the organization, development, and oversight of the overall Census 2000 QA program in the Census Bureau.

As mentioned earlier, we were to review these specific field activities:

- | | |
|--------------------|---------------------------------|
| - Block Canvassing | - Nonresponse Followup |
| - Update/Enumerate | - Coverage Improvement Followup |
| - Update/Leave | - LUCA 1998 Field Verification |
| - List/Enumerate | - LUCA 1999 Field Verification |

The listed activities involve enumerators receiving training on performing specific operations, collecting information, making entries on maps, and maintaining control lists and related forms.

The Census Bureau has not yet completed its evaluations of the effectiveness of its field QA approaches, which should provide information on workers' performance during production, the type and magnitude of production errors, and workers' perceptions of QA. Its review also is expected to identify deficiencies in the QA process, such as the lack of integration between production and QA activities and the poor implementation of some of the QA programs. Lacking such information, assessment of both the strengths and weaknesses is somewhat incomplete. Nonetheless, our discussion will focus on the broader aspect of the strengths and weaknesses of the QA program implemented for the data collection phase, reflecting both the views and opinions provided us and our review of assorted materials. We also note that a given aspect might be seen by some as a "strength" and, by others, as a "weakness." Our summary of the "strengths" and "weaknesses" of QA planning, organization, and implementation follows.

9.1

Strengths

- Consistent with its mission statement, whether in List Development or in Enumeration, Census 2000 continued the tradition, initiated in the 1960 Census, of incorporating into Census 2000 Field operations numerous activities described as QA. This commitment to quality and QA, demonstrated in five censuses over a 40-year period, certainly is a significant “Strength”;
- The objective for QA was that it be completely transparent in Census 2000 and, for the most part, it was. To that end, materials used to train enumerators and first level supervisors contained specific references as to why QA was important and to how it would be implemented, and all enumerators were exposed to the concept of and need for “quality” performance and, accordingly, measured against the established standards;
- Based on the perceptions of a diverse number of participants in Census 2000 and the on-going evaluation of the Census results, the QA activities are seen as “broadly” successful: they provided first level supervisors with relatively “real time” information on the quality of the enumerators, on their knowledge of how to carry out the activity, on the quality of their work and, to a lesser extent, on the quality of the information collected;
- The overall perception throughout the Bureau, and at all levels, even given the extent of negative comments, is that the Census 2000 QA Field program was an important element in preventing significant errors, and in preventing the clustering of significant errors. Although errors of both types did occur in selected instances, for the most part, they were caught expeditiously and rectified (the one glaring exception being the NRFU program, particularly emphasized by Hialeah). On this basis, the QA Field programs can be viewed, generally, as successfully meeting the first two elements of the Bureau’s QA mission: to prevent significant errors and to prevent the clustering of significant errors;
- Most operations, unlike the situation in the 1990 Census, had some form of QA process in place;
- The Census Bureau has committed itself to an extensive and comprehensive evaluation of all aspects of the Census 2000 program, which is still underway. Since the 1960 Census, the Bureau has been its own harshest and unstinting critic, conducting numerous evaluations and publishing extensive information on the quality of its programs and its results. These studies have been a valuable source, both for users and for improving future census methodology;

- Unlike the 1990 Census, Census 2000 is perceived by the broader public and the user community as “having been the most successful Census in many decades.” This sentiment is borne out by the fact that, unlike the period following the release of 1990 Census results, few if any local governments have contested the results of Census 2000 through legal action, or requested significant recounts; and
- Finally, Westat was asked to determine if the Bureau “made the best use of the available technology and statistical process tools with respect to our desire to promote timely and continuous improvement throughout the field operations.” In the context of the planning for the QA program for Census 2000, the answer is a definitive “Yes.” In dealing with what actually transpired, the perception is less clear, as discussed in previous sections and as detailed below.

9.2 Weaknesses

As a first point, we would note that a number of the so-called “strengths” also mask some “weaknesses.” Further, we have the benefit of hindsight and the luxury of time, both of which provide a broad frame in which to form any assessment:

- The fact that more than two years after the Census relatively little, if any, factual information is available currently concerning the effectiveness of the QA programs for the field activities noted above must be considered a major weakness. One would expect that the range of MIS reporting systems maintained to monitor field progress would have required and provided both on-going reporting of the efficacy of QA, and summary information on such detail as the number of interviewers fired because of QA failure, the number of assignments redone, summary information on the use of control charts or check sheets and the items found to have the most errors or, for that matter, on how widely and regularly those procedures in fact were used. Unfortunately, over two years after completion of the field effort, such is not the case. Neither does there appear to be debriefing materials or summaries, which could provide both insight and understanding of QA implementation in LCOs or RCCs. A lack of funding for the timely capture of the QA data undoubtedly was a major contributing factor, as was the absence of prompt and strong managerial support. We would strongly stress the importance of compiling, completing, and evaluating these activities in a very timely manner;
- Establishing the Deming QA philosophy as a goal for Census 2000 must be seen as a positive development. It remains unclear, however, as to whether the implications of such a decision were explored and, more importantly, fully documented, including an assessment of the Bureau’s ability to achieve it. Further, we were unable to locate documentation showing that an overall QA “plan” for Census 2000 was prepared or widely circulated, although the Bureau did hold a number of “Decennial QA Seminars, beginning in mid-1999, in order to acquaint staff with QA objectives, procedures, and responsibilities;
- Neither senior staff directly responsible for the QA effort nor the Executive Staff appear to have been closely involved in QA issues, including the scope, status or progress of the program. As best we could discern, QA issues rarely came to the

attention of the Executive Staff. The relatively low priority afforded QA at senior operating levels is reflected in the following: very limited QA research was undertaken during the period between the 1990 Census and Census 2000; evaluations of the 1990 Census programs (especially as related to the use of Quality Technicians in Field offices) do not appear to have been taken into account; a QA staff consisting of only six persons, some with very limited expertise in QA, was responsible for the QA planning for numerous major Field operations for Census 2000; and a lack of concern or emphasis on QA by the program Divisions resulted in QA activities being among the first items cut when budget stringencies arose;

- Relationships between those in the Decennial Statistical Studies Division (DSSD) nominally responsible for developing QA programs and the operating Division, in this case, the Field Division, on the whole, were best described as adversarial. Given its understandable focus on “production” and the necessity of completing the data collection task, the Field Division severely limited QA staff participation in the implementation of field operations, including observation, communication, information, or ability to modify. Leadership of the QA staff must share some of the responsibility for this state of affairs, since it failed to raise these issues to the attention of higher level census staff;

The problems brought on by the fact that production and quality responsibilities resided in different management areas were first highlighted in the 1980 assessment. As noted earlier, this issue is not unique to the U.S., having been mentioned and addressed by all three of the Countries contacted. Given this knowledge and history, getting top management in the respective areas to work together, with an appreciation of the importance of the two goals, should be an essential requirement of Census 2010.

- There is general agreement that funding for the planning and testing of QA programs for Census 2000 was insufficient and that few programs underwent adequate testing. More importantly, a vital aspect of the QA program—real-time capture and dissemination of QA data—was not implemented because of insufficient funding;
- Involvement in developing QA programs and materials is not seen as “having any cachet” in the Census Bureau. Personal advancement for those in the QA area is seen as seriously limited, and opportunities elsewhere in the Bureau are considered circumscribed. The same appears to be true of the QA programs themselves; which do not appear to be viewed as important or priority aspects of the Bureau’s activities;
- As in past censuses, the Bureau appears again to have been hampered by the lack of sufficient time to perform its many required functions. QA, as one of the last activities, suffered particularly, especially so in the conduct of the NRFU, the largest, and one of the most important of the QA programs; and
- Outside of the initial inspection of a relatively small part of the workload, either in the field or in the Office, few of the QA programs gave, or could give, meaning at the time of data collection as to whether the data being collected were of acceptable quality. Similarly, the MIS systems, whether OCS2000 or Cost and Progress (C&P), were inadequate in their coverage of QA elements that would have permitted a “real-time” review of performance, or real-time capture of QA results which would have provided an indication of problems. In fact, a recent report on the OCS2000 quoted

Bureau staff as seeing the system as “designed to be a control system for field operations and, therefore, not intended to be used either as a status monitoring system or a management information system.”¹⁰ Further, QA staff was not permitted regular, necessary access to these data, given a mindset that “the damage had been done by the time the battle could have been fought over each problem.” We also would note that, to our understanding, QA data were not aggregated to uncover common errors, either within a given AA, within a given Office, or across LCOs.

The question remains as to whether the Census 2000 QA field program should be seen as a series of unrelated, independent QC programs, rather than as an integrated QA program, especially so in connection with the collection phases of the census. To paraphrase Deming, “quality results from the prevention of defectives through process improvement, not inspection, which judges the quality of finished products and scraps or reworks defective items.” Review of the training and field operating materials and discussions with diverse staff at many levels who had widely varying responsibilities in Census 2000 clearly leads to the conclusion that, in general, during the data collection stage of Census 2000, the Bureau did not succeed in implementing a QA program but, rather, carried out effective and timely QC programs. There was little to no process improvement and, to the contrary, significant inspection. In this regard, the Bureau did not meet the last goal in its QA mission, namely, “to promote continuous improvement.”

At the beginning of the project, the Census Bureau posed four questions that they hoped to have answered by this study. These questions are answered in some detail throughout the Report. At this point, we present the specific questions and provide a brief summarization of the replies:

1. (Q) The Bureau’s QA philosophy emphasized prevention. We screened applicants, tested trainees, gave practice fieldwork, and observed and tested workers at the beginning of their field assignments. What other preventive measures should we have considered before we permitted workers to work alone?

(A) We believe the actions taken by the Bureau were essential to and required for quality performance. However, a number of additional actions directed to this goal are found throughout the Report; among the more important, we note the suggestion, based on the experience of Statistics Canada, to have a group day of review and training, following immediately on the beginning of data collection. The Bureau also should establish an effective MIS, to provide staff very rapid feedback of any errors encountered in the early completed materials.

2. (Q) Did the Bureau make the best use of the available technology and statistical process tools with respect to its desire to promote timely and continuous improvement throughout the operation?

¹⁰ Census Bureau. (2002). Operations Control System 2000, System Requirements Study. Final Report. Prepared by Titan Systems Corporation. Census 2000 Evaluation. R 2.a. February 2002.

(A) Since our evaluation suggests that the Bureau was not successful in promoting “timely and continuous improvement throughout the operation”, the response is negative. Although the Bureau’s planning assumed a “best use” approach, it fell short in implementing the QA programs at all levels of the effort, ranging from, among others, failure to establish a high level QA coordinating group, to the absence of real-time measures of quality, to allowing interneccine argument to prevent the timely review and correction of QA problems.

3. (Q) What limitations should the Bureau have taken into account when it adapted Deming’s management philosophy for its field operations?

(A) Most importantly, the complex nature of the undertaking, which renders it different from any example given (or possibly contemplated) by Deming. Specifically, many have commented on the impossibly short time period in which a decennial census must be completed, and how such an impossible time schedule works against the Deming concept of “continuous improvement” during the Field period. The vast scale of its operations is another factor mitigating against successful adoption, much less implementation of the Deming principles. Finally, the Bureau faced combining both of the foregoing factors with a virtually new, impossibly large, and wholly unskilled workforce. Given such obstacles, the Bureau is to be congratulated on its QA accomplishments.

4. (Q) What important new developments in the field of quality should the Bureau consider for its 2010 program?

(A) We have been unsuccessful in identifying any “important new developments”. We have mentioned throughout a number of actions which we believe the Bureau, should take to improve its QA programs. At this point, we would highlight the need for the Bureau to establish close relationships with countries such as Canada, Australia, and the UK to keep abreast of their developments in this area, to host a conference early in the decade to determine what new developments exist in the private sector and/or other government sectors, and to foster in-house research as a means of finding new approaches to QA. The final chapter of this report discusses additional recommendations for Census 2010 QA.

10. PLANNING QUALITY ASSURANCE FOR CENSUS 2010

The purpose of this section is: (1) to propose actions and approaches which address major shortcomings identified in the QA process for Census 2000, and (2) to offer creative solutions that are workable for a large and diverse workforce, taking into account both existing technology and technology that may be available in the next decade.

10.1 Rectifying Major Shortcomings

The Census Bureau currently is in process of producing a number of QA Profiles that will provide further insight into and knowledge about the effectiveness of the QA programs in identifying problems and enhancing quality. At the moment, virtually all these efforts are somewhat behind schedule. We believe that the expeditious completion of these profiles is an essential first step in the planning for Census 2010. Coupled with the on-going evaluation program of the Bureau's operations during the recent decennial census, these reports should identify the Field programs or operations that had significant problems, as well as those specific QA programs which were deemed to have failed or encountered difficulties. Using this information as a guide, the Bureau should develop alternative approaches that rectify the known deficiencies, through suggestions for program modification or restructuring. Such an exercise would provide a 'head start' in developing similar operations for Census 2010 and in avoiding the identified pitfalls. In addition, it would develop both an awareness of the types of problems that may be encountered, and possible solutions. Building on the past may not always be possible or even desirable; simply ignoring the past, however, is always foolish.

Concurrently, and with a sense of some urgency and priority, the Bureau also should address the issue of QA versus QC in a decennial activity. For the most part, Bureau staff has indicated that, in the true sense, QA did not exist in all phases of Census 2000; rather, that the QA programs carried out in the Field were, in fact, QC programs, that is, they called for significant inspection and, hopefully, repair, but provided little continuous feedback during the process. Others have maintained that QA was a reality during early planning and testing, but that QC is all that one can expect during the data collection phase, for all the reasons enumerated earlier in Section 7. We conclude that this issue, in reality, is but a "straw man," in other words, irrelevant to the Bureau's goal of ensuring the highest possible quality in decennial results. Nonetheless, it seems both to occupy and concern. We see no reason why one or the other, or both—QA and/or QC—cannot be used where they are most efficient and serve the common goals of improving quality and meeting the objectives of the QA program. It remains incumbent upon the Bureau, however, to address the issue of attempting to provide relevant information concerning the data collection

process on a timely schedule which allows the data to be useful in monitoring and improving quality. Such a QA step is the only method for truly improving quality during a census field period. A number of different approaches should be developed and tested, including an effective MIS, predesignated samples, flow processing, and greater emphasis on integrating quality into the operation, as examples, thus permitting the Bureau to meet the final goal of QA, namely, “to promote continuous improvement.”

An example of how QA might be used in “real-time” during the field operation would require establishing acceptable thresholds for measured variables. During the data collection stage, supervisory staff regularly monitored a series of variables collected in each LCO, including vacancy rates, the number of single person households or those containing large numbers of residents, etc. The goal was to detect outliers or those LCOs that seemed out of line with expectations. For the most part, the review was subjective in nature, using the extensive personal knowledge of the supervisory staff. Little or no evaluation was undertaken on a real-time basis to determine if the process could be improved.

At a minimum, cut-off levels could have been established in advance, utilizing the results of pretests or other types of available information. The levels could be variable, depending on the characteristics of the LCO, but consistent in establishing the point beyond which the results required investigation. An LCO exceeding established limits (since both minimum and maximum levels might be set) would be expected to review and explain the discrepancy. An automated procedure would identify the suspect LCOs, and the results of the reviews would be reported in real-time. In such a scenario, cut-off levels could be adjusted upwards rapidly if few errors were detected (thus saving time, effort, and funds) or, conversely, interviewers and other staff could be notified quickly to modify their behaviors if large numbers of LCOs were observed to be failing the established standard. Further, a sample of LCOs that fell just below the established threshold also could be reviewed to determine if the cut-offs were too permissive and allowing errors to slip through the system. Utilizing this type of approach, staff would both learn from the on-going activities and be able to adjust the level of review to ensure achieving the desired quality standard.

Preferably, control charts could be used to monitor the variables. Initial cut-off levels would be based, as above on earlier test results or expert knowledge. Using the observed distributions reported from LCOs and grouped on the basis of the designated variable, it would be possible to modify the cut-off levels on an on-going basis. In addition to identifying outliers, this approach also would allow for the monitoring of “runs” and “trends”. For example, if an LCO consistently has above average rates but never exceeds the cut- off level, it could still be identified for further review. Further, control charts would be helpful in identifying any LCOs that might be “gaming” the data, that is, falsifying data to avoid exceeding the threshold.

Undoubtedly, many other examples of supervisory staff regularly reviewing data on the progress of the Census exist throughout the decennial census effort. These include not only the field activity, but also administrative data on employment, or computer systems, or costs. It is likely that many of these examples would lend themselves to QA approaches providing real-time results which would have the net effect of reducing costs, reducing burden on staff, and providing higher quality data more quickly.

Given the importance the Bureau has attached to this issue, we further suggest that it would prove useful to the Bureau to convene a conference of experts in QA, as early in the decade as feasible, in order to share its thinking, discuss the issue, and become aware of any important new developments in the field of quality that should be considered for the 2010 QA program.

The Bureau also must address and resolve the relationship problems between the QA staff and the operating divisions. QA staff must be an integral part of the process. They must be allowed the opportunity to participate in tests, to observe programs in action, and to be provided with information on a timely basis. They also must be included in the decision-making councils. Conversely, the concerns of the operating Divisions and the realities of the Census taking process must be recognized and reflected in the demands of the QA programs. Simply put, each of these groups must work as “partners” in the Census effort, not as “outsiders” or “antagonists.” An example of this dichotomy in Bureau thinking with regard to QA versus data collection is found in the OCS 2000 Requirement Study, which states that “NRFU was completed ahead of schedule” (see Background, page 1), although the NRFU Progress Reports showed reinterview production lagging far behind schedule. When the statement was questioned, we were informed that the QA reinterview phase of NRFU was not viewed as an integral part of the NRFU effort and, thus, was not taken into consideration in determining the completion date of the program. In fact, the reinterview operation was designed as the key QA element of the NRFU effort in order to provide a real-time measure of the quality of the interviewers’ work. For that reason, it was planned for completion 1 to 2 weeks ahead of NRFU closeout; in reality, the reinterview program continued well after NRFU production had reached the 100 percent mark and was reported as “finished.”

To carry this point somewhat further and more pointedly, we believe that this burden falls directly on the Executive Staff and the Field operation. The Executive Staff must unambiguously mandate such cooperation and must routinely verify it is occurring. It then becomes incumbent on senior staff to see that this message is carried, understood, and acted on throughout its entire organization; that is, from the Associate Director for Field Operations, through the leadership of the Field Division, through the Regional Offices, through the RCCs, to the LCOs, and to every one of the staff at each of these levels. This message must state clearly and emphatically that QA is an integral and required responsibility of the entire Field organization, from Headquarters to RCC to LCO. All staff are responsible for both the

production, and the quality, of their data. The QA staff who come to observe or comment are to be seen as contributing to the work of the Field, not hindering its effort. Finally, there must be full awareness that failure to cooperate will not be acceptable and will be dealt with swiftly. Failure to disseminate such a message early in the decennial cycle and to ensure its implementation will surely prolong what has been a continuing problem for all too long.

To address these and other concerns, we propose that, at a very early point in the planning process for Census 2010, the Bureau:

- Establish, publish, and disseminate its QA goals and objectives. Dissemination should be to every individual involved in the decennial process, and prominently noted in training, observation, and review;
- Assign, implement, and monitor responsibility for QA. How best to accomplish such a recommendation is left to the Bureau, but we would suggest, as an initial step, the active and continuous participation of the Associate Director for Methodology and Standards, or someone on the immediate staff of the Associate Director, followed by QA representation in all participating operating entities, including Field. In this context, the Bureau may wish to explore the experience in the UK, which established a QA Czar, who had executive authority in the QA area;
- Integrate QA into the planning and implementation process by providing QA “a seat at the decision table”;
- Establish an experienced QA staff and provide it with a clear mandate as to its authority and responsibilities, especially in its interaction with the operating divisions;
- Ensure that the QA program is adequately funded, commensurate with its responsibilities and, further, is treated as an equal in budget reviews, rather than as “last in, first out”; and
- Mandate QA as an integral component of every facet of planning for Census 2010.

The Census Bureau also should implement a program to develop, test, and evaluate suggested QA programs and, further, should require that testing of proposed operational programs include the appropriate QA elements, thus allowing both the proposed programs and the QA measures to be evaluated and refined as a single whole, rather than as separate and disparate pieces. Such an approach presupposes a Bureau commitment to a sufficient cadre of experienced staff and funding adequate to the task.

At this point, we would note several suggestions incorporated in an internal Bureau report, which summarize much of the foregoing.¹¹ Specifically:

- Implement QA programs across all divisions;
- Communicate the reinterview QA operations purpose and expectations clearly to all field staff;
- The NRFU reinterview program should be tested during the decade;
- To be effective, the NRFU reinterview operation needs to be conducted on schedule and as planned; and
- Develop a system for real-time, up-to-the-minute data entry.

In reviewing QA in earlier censuses, we noted the use of a QA Technician Program, most recently as part of the 1990 Census, which included placing a trained technician in each of the 13 Regional Census Centers. The objectives of the RCCs QA Technician program were to promote management awareness of the purpose and importance of the various quality programs and to monitor adherence to the QA procedures. Although the Bureau's evaluation of the Census 1990 program concluded that "the QA Technician Program accomplished all of its objectives, in general," and recommended changes to make the program more effective in future censuses, this program was not included in Census 2000; we were unable to determine the reasons for this omission. Despite this decision, however, and given what we infer to have been a lack of knowledge and concern at the RCC level about QA programs and their importance (and, perhaps, even at the LCO managerial levels), we would strongly recommend reconsideration of this program for Census 2010. We also support the suggestions in the Census 1990 evaluation to establish a full-time position in each regional census center, that the persons selected to fill the positions be identified early in the census cycle, and that the extent of statistical training or expertise required for these positions be explored and established.

A key element of any successful QA program is its ability to provide an early warning of problems, so they can be identified, addressed, and action taken before the bulk of the activity has been completed. In a decennial census setting, this requirement has proven difficult to accomplish for most activities, if not all. One approach to be considered is the establishment of an early warning system, which would provide an indication of widely shared problems. For example, trained current survey interviewers could be used to prepare an advance listing of addresses contained in a national sample of AAs. As prelist operations are completed, this sample of AAs could be sent to a processing office for quick comparison with the advance listings. Such a sample check would provide a quick, efficient, and independent means

¹¹ Census Bureau. (2002). *Assessment Report for Nonresponse Followup*. Final Report. Census 2000 Information Memorandum #127.

of measuring the quality of the prelist operation and act as an early indicator of any problems in instructions, forms, or processes. Similarly, specially selected samples could be used to provide quick and inexpensive measures of mail response and vacancy rates, as well as of problems with the mailing or the content. Although not applicable to all field activities, it could be used where feasible and prove most helpful in assessing quality quickly.

We also are aware that web-based data collection may prove more feasible for the 2010 Census, than in Census 2000. Even though this development is likely to account for a relatively small proportion of the total workload (the Bureau currently assumes between 15 and 25 percent of all households), the rapid accumulation of this kind of database may lend itself for use as a “test bed” for a quick evaluation of item nonresponse and selected measures of quality. We would stress the need to explore the development of other types of early warning measures that could be implemented at both the LCO and the RCC levels for the data collection phases of the Census.

At this relatively early point in the decade, a first vision of Census 2010 contemplates a much simpler task than that in Census 2000. Specifically, the Bureau assumes that only a short form will be required of all households, with long form information collected independently through the American Community Survey. Such an approach can be viewed as requiring a much smaller field staff, less training but permitting more emphasis on quality, less effort on the part of the field staff and the supervisory staff, perhaps even far fewer LCOs, and a much simpler processing environment.

In this same vein, the expected development of improved and faster technology may permit the Bureau to have completed forms sent on-line by Enumerators, as they are completed, to centralized sites for processing in a continuous flow of “unrelated” forms, rather than in batches corresponding to collection units (AAs), as at present. Such a development would have significant implications for quality, in that flow processing would be almost immediate and continuous, without the need to wait for the full completion of an AA before transmittal to a Processing Center. With flow processing, incoming forms could be sampled and reviewed sequentially and problems resolved and appropriate corrective instructions sent to Enumerators quickly and efficiently.

Assuming a technology that would allow data to be transmitted and entered more quickly, it would seem feasible to institute a MIS that collects and transmits more diversified detail, including information relevant to QA assessment, such as item nonresponse and edit failure, both more accurately and more quickly than in Census 2000. Data currently collected as part of L/E, U/E, and NRFU such as number of vacant units, partial interviews, and persons per household could be included in a MIS that uses control charts to identify outliers and revise limits.

In discussing the scope and content of a proposed MIS, it is appropriate to reflect on and learn from the failure of the FMIS in the Canadian Census of 2001, which was seen as a nuisance and just ignored. The lesson, then, is to ensure that an MIS system meets the requirements of both providers as well as users; that is, it requests information that can be accessed easily, the system is easy to operate, it is flexible in application, and it provides users with what they need when they need it. In addition, both providers and users understand the need for and importance of the system and have received adequate training in using the system.

An optimal system should provide, in real-time and in simple detail, information to those who should know and care that the QA/QC activities are working and effective.

At a somewhat more detailed level, we would note that the Census Bureau is a vast repository of decennial experience, resident in the many staff members who participated in one way or another in Census 2000, and covering all its many programs and at all levels of implementation. At the time we began our investigation, we found that little of this experience had been systematically recorded, summarized, synthesized, or disseminated. We are most pleased to note that, at the current moment, the Bureau has begun the process of issuing such information as, for example, in Census 2000 Informational Memorandum #127, Assessment Report for Nonresponse Followup: Final, which was issued in September 2002. We would strongly urge the Census Bureau to continue to develop and produce such reports, even at this late date. We would note that the type of information that can be gleaned from such efforts could be both a valuable and important input into the planning for Census 2010. The many concerns voiced in the survey conducted among staff involved in the NRFU operation, as well as the issues noted in the report cited above, provide a valuable starting point.¹²

10.2 Other Suggestions for Consideration

Our exploration of census activities in several other countries did not reveal any particular approach or single program that could be said to provide the “key” to improved QA. In fact, the three countries studied, the UK, Canada, and Australia, are fully up-to-date on and knowledgeable about the latest approaches to census taking, including those in the U.S., and they adapted what others have done to fit their unique needs and requirements. It is both useful and worthwhile to highlight some of the facts, philosophies, and approaches, which merit consideration for Census 2010. Some of these suggestions have been noted earlier, but bear repeating because of the experiences encountered in other countries:

¹² Census Bureau. (2201). *Lessons Learned from the Census 2000 Nonresponse Followup Reinterview Operation: An analysis of questionnaires completed by a sample of the Reinterview Office Operations Supervisors, Telephone Clerks, and Enumerators*. Unpublished.

- The need for continuity of staff over time cannot be emphasized too strongly. Although obviously somewhat easier to accomplish in countries that use five-year cycles for their census programs, continuity is nonetheless an important requirement for the U.S. program, as well. Experience is an important element in understanding the diversity and scope in taking a census. At the same time, the selective addition of “new blood” can significantly invigorate the strength of any program;
- Given the size, scope, complexity, and pressures of a census, it is essential that staff work together collegially;
- Quality is likely improved by simplifying the Enumerators’ task. Removing, to the extent feasible, administrative responsibilities from the Enumerator and Crew Leader would seem a major step in accomplishing that objective;
- Statistics Canada believes that the one-day training session for census enumerators held immediately following its questionnaire drop off operation was an invaluable contributor to the quality of the census operation. Specifically, the session allowed enumerators to get immediate answers and feedback on problems encountered during their initial forays into their EAs immediately preceding the session, it provided supervisors with a single forum in which to address problems and concerns of all Enumerators, it offered a substantial boost to the morale of the Enumerators, and it provided immediate feedback to Headquarters of the common and most frequent problems and allowed them to provide immediate answers before the operation had progressed very far;

Even accepting the many differences between the Canadian approach and that of the U.S., we would suggest that this feature be explored and examined as a possible addition to Census 2010. Currently, Enumerators are visited by their CLs as soon as possible after the initiation of followup. However, the reality is that some of the enumerative staff is not visited until they are well into the followup operation. If, indeed, the 2010 Census requires the collection of only short form information, the significant decrease in complexity and workload throughout the operation, coupled with the need for a somewhat smaller field staff, may allow for a more flexible time schedule, in which an additional day of training would be both feasible and practical, as well as offering significant benefits in consistency, timing, and quality;

- Given the current scarcity of information concerning the effectiveness of the QA programs, we suggest that the Bureau examine the feasibility and value of undertaking an evaluation study similar to that carried out by Statistics Canada. The study utilizes an appropriate, pre-designated sample, which could consist of EAs, addresses, or both, depending on the design of the 2010 Census. Documents from the selected addresses, as well as control forms, as completed in the Field, are examined to determine whether the instructions given for completion were appropriate and appropriately carried out. In addition, these sample forms are used to prepare selected rates and measures which serve to evaluate the census field operations, such as coverage of households and persons, incoming error rates for selected questions, and rejection rates for the entire form. Such an approach should provide useful information in a timely manner, both to assist in evaluation the 2010 Census program, and in preparing for the next Census;

- Documentation—Experiences, problems, solutions, suggestions, and recommendations which rise to the surface during or following the hectic days of census taking are all too often forgotten or overlooked. The obvious solution is documentation and ready access to the information. Whether the approach followed in Canada, or Australia, or the UK is applicable, or even optimal, is not the issue. The fact remains that debriefing of staff at all levels, construction of an easily accessible data base containing such information, and supplemented with the results of memoranda detailing problems and issues and some containing solutions implemented at the moment, along with the suggestions of staff for improvements or changes, can provide a most valuable resource to those charged with responsibility for planning the next census. Senior management should be required to address these recommendations as part of planning the next Census. Finally, “documentation” is at the heart of Deming’s “continuous improvement;” and
- Since a number of countries will be undertaking censuses during the mid-decade period, the Bureau (including both Field and QA staff) should monitor the QA/QC approaches developed by these countries and determine if they might prove beneficial to Census 2010. In fact, the Bureau should seriously consider establishing an on-going relationship with the statistical agencies of the UK, Canada, and Australia, in order to “share and compare” past QA experiences, including quantitative results of their respective programs, and planning and program for future census QA activities, if not for all aspects of decennial planning.

More broadly, we strongly urge the Bureau to ensure that these fundamental requirements of the Deming philosophy become an integral part of the Bureau’s philosophy:

- QA must be part of management; not a single nor a separate operation; and
- QA must be part of the management responsibility at the operating level, including not only support, but also both awareness and monitoring.

Finally, perhaps the Deming application in a census setting must be understood and seen as a “process improvement over time,” one in which the experience of one Census informs the next.

REFERENCES

Block Canvassing

Crew Leader's Manual Block Canvassing—August 1998

Crew Leader's Manual Supplement Block Canvassing—November 1998

Block Canvassing Lister's Instructions—January 1999

Coverage Improvement Followup

Coverage Improvement Followup Crew Leader Manual—March 2000

Daily Supervision

Performing Quality Assurance Dependent Check (continued)—October 1999

List/Enumerate

List/Enumerate Crew Leader Manual—October 1999

LUCA 1998

Luca 1998 Field Verification Supervisor's Manual—April 1999

Luca 1998 Field Verification Lister's Instructions—May 1999

LUCA 1999

Luca 1999 Field Verification Lister's Instructions—N/A

Luca 1999 Field Verification Supervisor's Manual—April 1999

Nonresponse Followup (NRFU)

Nonresponse Followup Enumerator Manual—July 1999

Update/Enumerate

Update/Enumerate Enumerator's Manual—September 1999

Update/Enumerate Crew Leader's Manual—October 1999

REFERENCES (CONTINUED)

Update/Leave

Update/Leave Enumerator's Manual—July 1999

Update/Leave Crew Leaders Manual—August 1999

Update/Leave Office Review Exercise Answer Key—September 1999

Reinterview

Office Operations Supervisor (OOS) for Reinterview Training Checklist—N/A

Field Operations Manual—September 1999

Reinterview Training Guide—October 1999

Field Operations Manual—January 2000

General

Census Bureau. (2002). *Assessment Report For Nonresponse Followup: Final Census 2000 Informational Memorandum #127*, [Unpublished].

Census Bureau. (2000). *Nonresponse Followup, Program Master Plan: Revision 1 Census 2000 Informational Memorandum #26, R-1*, [Unpublished].

Census Bureau. (1999). *Census 2000 Operational Plan* [Unpublished]. Report: Updated Summary Document.

Census Bureau. (1999). *Decennial QA Seminar* [Unpublished]. Document Prepared for Internal Briefing.

Census Bureau. *Master Address File (MAF) Building Operations* [Unpublished]. Document Prepared for Internal Briefing.

Census Bureau. *DSSD Census 2000 Procedures and Operations Memoranda* [Unpublished]. Documents: Selected Series and Numbers.

Census Bureau. *Selected Trip Reports* [Unpublished].

REFERENCES (CONTINUED)

Publications

- Australian Bureau of Statistics. (2000). *How Australia Takes a Census*. 2001 Census of Population and Housing.
- Census Bureau. (2002). *Management Information System 2000, System Requirements Study*. Census 2000 Evaluation R.3.c.
- Census Bureau. (2002). *Operations Control System 2000, System Requirements Study*. Census 2000 Evaluation R.2.a.
- Census Bureau. (1993). *Effectiveness of Quality Assurance: 1990*, Series CPH-E2.
- Census Bureau. (1976). *U.S. Census of Population and Housing: 1970*, Procedural History. PHC ®-1.
- Census Bureau. (1973). *1970 Census of Population and Housing*, Procedural History: Advance Issuance of Chapters 13 and 14. PHC(R)-1B.
- Census Bureau. (1967). *United States Censuses of Population and Housing 1960*, Quality Control of the Field Enumeration.
- Corey, Stephen R. (1990). *The Seven Habits of Highly Effective People*. Fireside Book. Simon & Schuster, Inc.
- Deming, W. Edwards. (1993). *The New Economics for Industry, Government, Education*. MIT Center for Advanced Engineering Studies.
- Deming, W. Edwards. (1982). *Quality, Productivity and Competitive Positions*. MIT Center for Advanced Engineering Studies.
- Deming, W. Edwards. (1982). *Out of the Crisis*. Massachusetts Institute of Technology. Cambridge.
- Dillman, Don A. (1996). Why Innovation is Difficult in Government Surveys. *Journal of Official Statistics*, Vol. 12, pp. 191-197.
- Morganstein, D.R., and Hansen, M.H. (1990). *Survey Operations Processes: The Key to Quality Improvement*. In Data Quality Control, Chapter 8. Marcel Dekker, Inc. New York.
- Statistics Canada. (2000). 2001 Census, Collection.
- Statistics Canada. (Undated). *2001 Census of Canada, Quality Control Technician's Manual*, Form 70.
- Statistics Canada. (2002). *1996 Census, Evaluation of the Field Quality Control*.

RESPONSE TO WESTAT FINDINGS BY DECENNIAL STATISTICAL STUDIES AND FIELD DIVISIONS

The Census Bureau commissioned Westat Corporation to produce this report, "Evaluation of the Census 2000 Quality Assurance Philosophy and Approach used in the Address List Development and Enumeration operations" to critique its decennial quality assurance (QA) program. Due to time and budget constraints, the Census Bureau asked Westat to focus on the QA program used in the field operations that updated the Master Address File (MAF) and TIGER system, and the field operations in which personal visit data collection was conducted. In their reviews of this evaluation report, several Census Bureau divisions raised concerns about the methodology that Westat had employed in compiling in this report and its reliance upon interviews of a relatively small number of Census Bureau Headquarters staff. In some cases, the summary of views implies situations or policies that weren't necessarily true. In their report, Westat makes a number of points that unless placed in the proper context may mislead the reader. In this response, we address these points in general and attempt to place them in the proper context and setting in which they occurred during Census 2000.

This response is prepared in full recognition that while Census 2000 was unquestionably the most successful Census taken, it was far from perfect and there are a number of important lessons to be learned and areas in which improvement can be made, including in the design and implementation of essential QA programs and inter-divisional working relationships.

As the Census Bureau conducts its early planning, research and development for the 2010 Census, the Decennial Statistical Studies Division and the Field Division, working with the Decennial Management Division have prepared a Memorandum of Understanding which will set the framework for more cooperative working relationships in the development of quality assurance and quality control programs for the address list enhancement and field data collection operations that will be included in the 2010 Census. Based upon the Census 2000 lessons learned and feedback from debriefings of Regional Census Center and Local Census Office staff, the management structure of the 2010 Census of both the Regional Census Center and the Local Census Office are being expanded to include management positions and as needed support staff dedicated to QA. This expanded management structure will be implemented, and refined as necessary, in the 2004 Census Test, the 2006 Census Test, the Dress Rehearsal, as well as, the 2010 Census.

This response addresses the following concerns about this Westat report: (1) Westat's finding that the Census Bureau did not have a single, comprehensive QA plan, but instead separate QA programs for each field operation; (2) Westat's assertion that there was a lack of a senior management team for coordinating and approving the overall QA plan and reviewing the implementation; (3) The lack of current QA results data and the implications on management of the program; and (4) The perception of the QA program based on Westat interviews.

Comprehensive QA Plan

The Westat report does not fully reflect the complex and demanding context in which Census 2000 was conducted. Managers of Census 2000 operations, both at Census Headquarters and working in the field offices had to balance quality assurance with meeting tight deadlines, staffing large-scale, people-intensive operations during a period of historically low levels of unemployment, and close monitoring of costs and expenditures. Due to factors not under the Census Bureau's control, such as the January 1999 Supreme Court decision, the operational design of Census 2000 placed unprecedented deadline pressure upon the completion of all field operations on schedule. Completing the largest single Census 2000 field

operation, Nonresponse Followup, on schedule was absolutely essential to deliver the apportionment counts to the President by December 31, 2000, as mandated by public law. Census 2000 was also conducted and completed under intense internal and external scrutiny. Virtually all details and activities of Census 2000 were followed and assessed by the Department of Commerce Inspector General Office, the General Accounting Office and the Census Bureau Monitoring Board.

Before the QA staff developed the individual Census 2000 QA programs, the QA manager briefed them on the history of QA in decennial operations, provided them a booklet containing profiles of the QA operations used in the 1990 census, and informed them of the 2000 QA philosophy and approach. In fact, in June of 1999, the QA staff gave a seminar to inform their colleagues in the Decennial Statistical Studies Division about the decennial QA philosophy and approach. The seminar provided an overview of the individual QA programs, but more importantly showed how these programs, although seemingly different, all had the same underlying philosophy and approach: prevention, improvement, and protection. Essentially, a general plan and approach for QA in Census 2000 existed, but was not formally documented prior to development of individual census operations.

Senior Management Involvement

Our Census 2000 QA philosophy and approach stemmed from research of the literature and lessons learned in previous QA programs. Development of the individual QA programs was a team effort. QA staff working with members of an operation's planning team developed the QA program for the operation. The QA plan for each operation was provided in the operation's Program Master Plan (PMP) - a high level document (reviewed and approved by senior management) that outlined the operation from start to finish. Senior management established a QA staff in the decennial area to manage and carry out the QA plan. The team approach demanded different methods of managing the program. Supervisors of team members advised their staff regarding the substance of the program rather than actually managing the operational development of the program. The teamwork approach involved in developing programs did not lend itself towards giving (or having the appearance of giving) senior management direct guidance of a consistent and comprehensive QA program for Census 2000.

However, senior management and executive staff supported and monitored progress of the QA program throughout the census. In addition, senior management maintained an independent QA staff within the Decennial Statistical Studies Division responsible for developing and overseeing the QA program. During Census 2000, the senior management team occasionally had to make difficult decisions to balance the needs of QA and census operations. For successful completion of the census, the management approach must include the flexibility to make such decisions.

Statements about the commitment to and conduct of QA by the Regional Census Centers and the Local Census Offices are included in the Westat report, but these statements do not represent the views of the regional staff since Westat did not conduct interviews with Regional Directors or Assistant Regional Census Managers.

Current QA Results Data

The results of the QA operations were not captured in real-time. The Census 2000 Management Information Systems did provide timely data on costs and progress (e.g., cases assigned, cases completed), but the results data such as pass/fail rates were not available during the census. A consequence of not having results data in real-time was that senior management lacked the tools they needed to monitor and ensure that the quality assurance operations were being completely carried out. However, the QA manager brought QA issues to senior management's attention for corrective action. In

deciding upon appropriate corrective actions, senior management was responsible for resolving quality issues while at the same time maintaining critical progress towards completion.

Senior management did have cost and progress data at their disposal and they did follow the progress of QA to see the degree to which the job was completed. The local census offices had results information available to them since the paper forms on which results were recorded were submitted to the offices prior to being sent for processing. These forms provided the local census offices with the ability to identify areas which needed special attention to determine appropriate actions (e.g., re-check an enumerator's work, retrain an enumerator, or release an enumerator from the operation).

Perceptions of the QA Program

Westat was asked to proceed with this evaluation before the Census Bureau's own evaluation of its QA program was complete. The individual QA profiles that make up this self-evaluation would have provided Westat objective information about our QA program. As a result, Westat had to rely heavily on information gathered through interviews with a small cross section of Census Bureau staff that were involved with Census 2000. Unfortunately, aside from one individual who worked at various management levels in a local census office, Westat did not interview any staff member who worked directly in the field operations or at the regional level.

The Westat report states that the elapsed time between completion of the Census and evaluation efforts made it impossible for them to conduct interviews with "representative sample" of the field staff. This statement suggests that knowledgeable field staff were no longer available for interviews. However, the Regional Directors had responsibility for conducting both the Census 2000 field operations and the associated QA programs, and the majority of those Regional Directors still occupy their permanent positions.

Because of these limitations, we must reiterate Westat's point that the opinions expressed in this report represent only the perspective of the individuals interviewed. These perspectives do not necessarily reflect truth. For example, if some interviewees perceived limited involvement by senior Census Bureau management, that perception could be flawed because interviewees may not have first-hand knowledge of the true involvement of senior management. Whereas the interviewee perspectives provide insight into the feelings and impressions of various levels of Census Bureau staff, they cannot be taken to necessarily reveal factual information regarding how the census operations and associated QA were conducted.

In closing, the objective of this response was to provide a background for a number of points made in this report so that the reader could place them in the proper context. We've reiterated the point that the opinions expressed in this report are those of a relatively small number of people and may not necessarily reflect truth, but merely perception. Overall though, their recommendations are clearly aimed at providing the Census Bureau with guidance for taking the weaknesses that existed in 2000 and using them to develop and implement improvements for future censuses. We have already begun this process for our census testing activities leading up to the 2010 Census.

APPENDIX A

Contributing Participants

U.S. Census Bureau

Martin Appel, Statistical Research Division

Cynthia Clark, Associate Director for Methodology and Standards

Howard Hogan, Chief, Decennial Statistical Studies Division

Carrie Johanson, Decennial Statistical Studies Division

Ruth Ann Killion, Chief, Planning, Research, and Evaluation Division

Gail Leithauser, Assistant Chief, Field Division

Broderick Oliver, Decennial Statistical Studies Division

Rebecca Piegari, Decennial Statistical Studies Division

Jennifer Reichert, Decennial Statistical Studies Division

Gabriel Sanchez, Field Division

Jimmie Scott, Demographic Statistical Methods Division

Peter Sefton, Field Division

Carol Van Horn, Assistant to the Associate Director for Field Programs

Preston J. Waite, Associate Director for Decennial Census

David Whitford, Decennial Statistical Studies Division

Office for National Statistics, United Kingdom

Rod Massingham, Head of Data Collection

Lesley Simmonds, Procedure & Instructions Manager

Andy Teague, Deputy Director of Census

Statistics Canada

Michael Bankier, Chief, Census Research and Development

Jean-Rene Boudreau, Senior Methodologist, Census Collection Methodology

Katherine McClean, Chief, Address Register & Geography Methods

Mike Sheridan, Assistant Chief Statistician, Social, Institutions and Labour Statistics Field

Australian Bureau of Statistics

John Struik, Head, Census Programs

Paul Williams, Head, Development of Field Operations

APPENDIX B

Deming's 14 Management Points

1. Create constancy of purpose toward improvement of product and service, with the aim to become competitive and to stay in business, and to provide jobs;
2. Adopt the new philosophy. We are in a new economic age. Western management must awaken to the challenge, must learn their responsibilities, and take on leadership for change;
3. Cease dependence on inspection to achieve quality. Eliminate the need for inspection on a mass basis by building quality into the product in the first place;
4. End the practice of awarding business on the basis of price tag. Instead, minimize total cost. Move toward a single supplier for any one item, on a long-term relationship of loyalty and trust;
5. Improve constantly and forever the system of production and service, to improve quality and productivity, and thus constantly decrease costs;
6. Institute training on the job;
7. Institute leadership. The aim of supervision should be to help people and machines and gadgets to do a better job. Supervision of management is in need of overhaul, as well a supervision of production workers;
8. Drive out fear, so that everyone may work effectively for the company;
9. Break down barriers between departments. People in research, design, sales and production must work as a team, to foresee problems of production and in use that may be encountered with the product or service;
10. Eliminate slogans, exhortations, and targets for the work force asking for zero defects and new levels of productivity. Such exhortations only create adversarial relationships, as the bulk of the causes of low quality and low productivity belong to the system and thus lie beyond the power of the work force;
- 11a. Eliminate work standards (quotas) on the factory floor. Substitute leadership;
- 11b. Eliminate management by objective. Eliminate management by numbers, numerical goals. Substitute leadership;
- 12a. Remove barriers that rob the hourly worker of his right to pride of workmanship. The responsibility of supervisors must be changed from sheer numbers to quality;
- 12b. Remove barriers that rob people in management and in engineering of their right to pride of workmanship. This means, inter alia, abolishment of the annual or merit rating and of management by objective;
13. Institute a vigorous program of education and self-improvement; and
14. Put everybody in the company to work to accomplish the transformation. The transformation is everybody's job.